Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *University of Michigan*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Reaction Time and EEG Activation Under Alerted and Nonalerted Conditions: R. W. LANSING, E. SCHWARTZ, AND D. B. LINDSLEY 1
Performance in Differential Conditioning and Discrimination Learning as a Function of Hunger and Relative Response Frequency: K. W. SPENCE, K. P. GOODRICH, AND L. E. ROSS 8
On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall: J. DEESE 17
Absolute Judgments of Odor Intensity: T. ENGEN AND C. PFAFFMANN 23
The Relationship Between Probability Difference (π_1-π_2), and Learning Rate in a Contingent Partial Reinforcement Situation: P. J. WOODS 27
Aftereffect of a Moving Pattern: V. R. CARLSON 31
Motion Parallax as a Determinant of Perceived Depth: E. J. GIBSON, J. J. GIBSON, O. W. SMITH, AND H. FLOCK 40
A Comparison of Three Varieties of Training in Human Problem Solving: L. MORRISETT, JR. AND C. I. HOVLAND 52
Reversal and Nonreversal Shifts in Kindergarten Children: T. S. KENDLER AND H. H. KENDLER 56
Effect of Magnitude of Reinforcement on Acquisition and Extinction of a Running Response: H. L. ARMUS 61
An Evaluation of the Manifest Anxiety Scale by the Use of Electromyography: A. M. ROSSI.. 64
Response Learning in Paired-Associate Lists as a Function of Intralist Similarity: B. J. UNDERWOOD, W. N. RUNQUIST, AND R. W. SCHULZ 70
The Effect of Precision, Delay, and Schedule of Knowledge of Results on Performance: F. J. MCGUIGAN 79
The Effect of Differential Onset Time on the Conditioned Response Strength to Elements of a Stimulus Complex: D. D. WICKENS, R. S. GEHMAN, AND S. N. SULLIVAN 85
Supplementary Report: Acquisition, Extinction, and Counterconditioning With Different Verbal Reinforcement Combinations: E. L. FERGUSON AND A. H. BUSS 94

American Psychological Association
Vol. 58, No. 1 July 1959

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.*** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1959 the publication cost is $30 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1959 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 58, NO. 1 JULY, 1959

REACTION TIME AND EEG ACTIVATION UNDER ALERTED AND NONALERTED CONDITIONS[1]

ROBERT W. LANSING,[2] EDWARD SCHWARTZ,[3] AND DONALD B. LINDSLEY

University of California, Los Angeles

Berger (1930, 1932) first observed that visual stimulation suppressed or blocked the alpha waves of the human electroencephalogram (EEG). Since any type of sensory stimulation produced an alpha blockade he interpreted the effect as a generalized cortical inhibition, after Pavlov. Subsequently, he held that anything which tended to focus attention upon a stimulus or task blocked the alpha rhythm over all areas of the brain. Adrian and Matthews (1934) felt that some form or pattern in the visual field was necessary to block the occipital alpha rhythm, but agreed with Berger that the focusing of attention, rather than the specific type of sensory stimulus, had a widespread effect upon the alpha rhythm. Other studies dealing with the effect of sensory stimulation and attentional states have been reviewed by Lindsley (1944) and Ellingson (1956).

[1] Supported by USPHS Grant B-362, Division of Neurological Diseases and Blindness, National Institutes of Health.

[2] USPHS Postdoctoral Fellow during course of this research; now at University of Texas.

[3] Now at Veterans Administration, Chicago, Illinois.

Thus alpha blockade came to be associated with sensory stimulation of various types and was found over various areas of the brain irrespective of the type of stimulation. Furthermore, startle, anticipation, or the assumption of an attitude of attention or problem set, also clearly interrupts the on-going alpha activity and tends briefly to hold it in abeyance. Brief sensory stimulation, such as a flash of light, generally produces only a short interruption of the alpha rhythm, perhaps 1 sec. in duration. Unexpected and slightly startling stimuli often result in suppression of alpha waves for a longer period. Attempts to maintain a focus of attention on a visual target, or attempts to track a moving target with eye and hand usually disrupt the alpha rhythm for some time until the task becomes habitual and somewhat automatized. Concentrating upon a problem in mental arithmetic often suppresses the alpha rhythm and holds it in abeyance until an answer is given.

From the above observations of generalized alpha blockade in the EEG in association with various psychological phenomena it became

apparent that some diffusely operating influence was making itself felt upon the electrical state of the brain. The nature of this process and its mechanism remained obscure until Moruzzi and Magoun (1949), and others (Lindsley, Bowden, & Magoun, 1949; Lindsley, Schreiner, Knowles, & Magoun, 1950), demonstrated that both the electrocortical activity and the behavior of animals could be strikingly modified by electrical stimulation of the reticular formation, or by interruption of the reticular core of the lower brain stem. The upward projections of the reticular formation have come to be known as the *ascending reticular activating system* (ARAS).

Electrical stimulation shifted the patterns of electrical activity at the cortex from a sleeping or awake resting state with synchronized or organized alpha rhythms to a picture of low amplitude, fast waves, with alpha or other synchronized waves absent. Lesion or interruption of the ARAS in the midbrain tegmental region produced synchronized slow waves and spindle bursts as seen in sleep and the behavior of the animal was characterized by sleep or coma. On the other hand, lesion of the specific sensory pathways (medial and lateral lemniscii), leaving the reticular formation intact, had no such effect; the EEG and the behavior were those of a normal waking animal.

It was demonstrated that any type of sensory stimulation produced behavioral arousal or alerting of the animal and that when it did so the electrocortical activity changed to an "activation" or "desynchronization" pattern (Lindsley et al., 1950; Rheinberger & Jasper, 1937). Such an effect is mediated by collaterals from specific sensory pathways (Magoun: 1952, 1954) in the lower brain stem which conduct impulses into the reticular network. The reticular formation has been shown to have widespread functional and anatomical connections with the cerebral cortex through diffuse thalamic and extrathalamic projections (Rossi & Zanchetti, 1957). More recently it has been shown (French, Hernandez-Peon, & Livingston, 1955) that stimulation of various areas of the cortex electrically will evoke responses in the reticular formation, thus extending the range of influence upon the reticular activating system beyond that of sensory influx. Such corticifugal connections appear to be capable of influencing the ARAS so that its effects upon the cortex and behavior reflect states of arousal, alerting, attention, or a problem-oriented set, as a result of past experience or special instructions (see Lindsley: 1956, 1958a, 1958b, 1958c).

Psychologists have long known that a warning or an alerting signal prior to a reaction stimulus reduces the reaction time (Woodworth & Schlosberg, 1954). The length and variability of the foreperiod interval determine its effectiveness in reducing RT (for recent studies see Klemmer, 1956; Teichner, 1954). Most of these investigators confined their observations to foreperiods ranging from 1 to 36 sec. and observed the greatest reductions in RT between 1 and 8 sec. Klemmer (1956) and Davis (1940) included foreperiod intervals below 1 sec. and each observed that minimal RT's occurred with foreperiod intervals of less than 1 sec.

Psychologically, therefore, set or attention is reflected in reduced RT. As mentioned above, the EEG alpha blockade is associated with sensory stimulation, attention and instructional set. Recent neurophysiological evidence, as indicated above, suggests that the ARAS underlies both electro-

cortical and behavioral effects of an alerting stimulus. Accordingly, therefore, our problem is to relate RT as a psychological or behavioral performance measure to EEG activation (alpha blockade) as a neurophysiological index of alertness or attention.

METHOD

Visual RT's were measured in nine normal adult *Ss* under two general conditions: *Nonalerted* and *Alerted*. During the nonalerted state the visual reaction stimulus was presented at times when the alpha rhythm of the EEG varied spontaneously and included patterns of activity characterized as (*a*) *good alpha waves*, i.e., strong, regular and rhythmic waves; (*b*) *poor alpha waves*, i.e., irregular or poorly organized waves of alpha duration; and (*c*) *no alpha waves*, i.e., times when alpha waves were absent. Under the alerted condition an auditory warning signal (buzzer) preceded the visual reaction stimulus by intervals varying from 50 to 1000 msec. Brief foreperiods of 50 to 250 msec. did not allow sufficient time for complete activation or alpha blockade to occur before the visual reaction stimulus was presented. Long foreperiods ranging from 250 to 1000 msec. permitted activation and alpha blocking to occur before the visual reaction stimulus. Figure 1 illustrates the sequence of events under the alerted condition for brief and long foreperiods.

The RT's were measured with a Berkeley Electronic Counter which was activated by 1000 per sec. pulses and counting was begun at the onset of the stimulus light and terminated when *S*'s finger depressed the reaction key. The light stimulus was a blue-white, 10 μsec. flash of 600,000 lumens peak intensity, produced by a Grass PS-1 Photostimulator. The *S* sat with eyes closed in a dark room 70 cm. in front of an opal glass screen 50-cm. square which was illuminated by the stimulus flash. The eyes-closed condition was used to better preserve or facilitate the alpha rhythm.

The EEG's were recorded from occipital and motor area leads (scalp to ear reference). Eye movements and blinks were recorded for control purposes. The recording instrument was a Grass Model III-D, 8-channel electroencephalograph, operated at a paper speed of 120 mm. per sec.

A preliminary practice session was given prior to the actual experiment. From 200 to

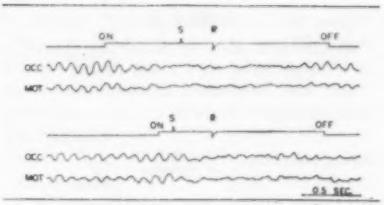

FIG. 1. Effect of EEG activation upon reaction time. Occipital and motor area alpha waves recorded during measurement of visual reaction time (S to R) under alerting produced by auditory stimulus (buzzer: on-off). Top record, long foreperiod, buzzer blocks alpha waves prior to onset of visual stimulus (S); reaction time short. Bottom record, foreperiod short, alpha waves not blocked before visual stimulus; reaction time long.

500 RT's were obtained during one or more experimental sessions. The RT series were broken up into groups of stimulations with brief rest periods interspersed. Under nonalert conditions *Ss* were instructed to press the key as quickly as possible after the light flashed, and then to relax until the next flash appeared. The stimuli were presented at intervals varying from 15 to 45 sec. during good, poor and no-alpha states as described above. Under the alerted conditions a buzzer warning signal preceded the visual stimulus at randomized intervals varying from 50 to 1000 msec. The *Ss* were told that the buzzer signal would precede each light stimulus by a varying interval and that they were to press the key as quickly as possible after the light flash.

RESULTS AND DISCUSSION

Table 1 shows the mean RT's obtained under nonalerted and alerted conditions. The over-all mean for the nonalerted state was 280 msec. and for the alerted state 225 msec. This difference is highly significant ($P < .01$ by Wilcoxon's matched-pair, nonparametric test). An even greater difference is noted between the over-all mean RT for the nonalerted condition and that for the alerted condition in which a long foreperiod (250–1000 msec.) and alpha blockade occurred

TABLE 1

MEANS AND *SD*'S OF REACTION TIMES FOR NINE *Ss* UNDER ALERT AND NONALERT CONDITIONS

Alerting Foreperiod	EEG at Time of Stimulus	Reaction Time	
		Mean	*SD*
Nonalerted	Good alpha	280	38
	Poor alpha	278	42
	No alpha	281	46
	All trials	280	
Alerted			
50–250 msec.	Alpha not blocked	255	42
300–1000 msec.	Alpha blocked	206	28
	All trials	225	

prior to the visual reaction stimulus. The lowest mean reaction time of 206 msec. was obtained under this latter condition.

There were no significant differences between the subconditions of the nonalerted state; both the means and *SD*'s were very similar. The alerted condition in which the foreperiod of warning was brief (50–250 msec.), and in which the alpha rhythm was not completely blocked by the time the visual reaction stimulus occurred, gave a mean RT of 255 msec. Although this shows some reduction in RT, due presumably to partial activation by the alerting stimulus, the difference between this mean and the over-all mean for the nonalerted condition is not statistically significant.

The variability as shown by the *SD*'s is similar under all conditions except that of the alerted state with a long foreperiod and alpha blocking prior to the visual reaction stimulus. The *SD* in this latter instance is markedly reduced from that of the other conditions and this difference is statistically significant.

The foregoing results thus confirm previous studies in showing that RT's may be significantly reduced by a preparatory signal. Furthermore the results show that minimal RT's may be obtained with foreperiods of less than 1 sec.; however, to be maximally effective the preparatory signal must occur at least 300 msec. prior to the visual reaction stimulus. The relation such an optimal foreperiod of alerting bears to the electroencephalogram will be taken up next.

There is shown in Fig. 2 the relaxed RT level at 280 msec. (dash line). The alerted RT curve shows a progressive reduction below this level as the foreperiod of warning increases from 50 to 300 msec. It may be seen that with foreperiods longer than 300–400 msec. very little further reduction in RT occurs. Preparatory intervals greater than 1 sec. were not studied since our primary interest was in the early phases of the alerting process as these are reflected in the EEG by by activation or alpha blockade.

The alpha blocking curve of Fig. 2 represents the percentage of trials for each foreperiod interval in which blocking was complete. It will be seen that the time course of this curve is almost identical to that of the

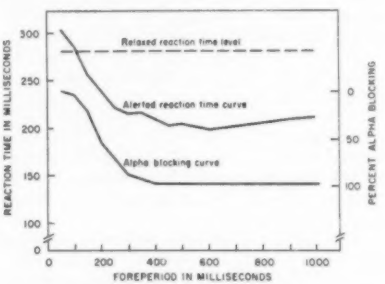

FIG. 2. Reaction time and alpha blocking plotted as a function of the foreperiod interval. Note that reaction time is reduced to a minimum when the foreperiod ranges from 300 to 1000 msec. and that the alpha blocking curve follows a similar time course.

alerted reaction time curve. In other words, as activation or alpha blockade becomes effective, RT's are progressively reduced. Alpha blocking time (latency from preparatory signal to alpha wave suppression or blockade) is typically 200–300 msec., as has been observed in many other studies employing a variety of sensory stimuli. The fact that the alerted reaction time curve and the alpha blocking curve each approach an asymptote at about the same point in time (300 msec.) strongly suggests that they are influenced or governed by a common process. This process appears to be the excitation of the ARAS, since activation or alpha blockade is typically produced by stimulation of the reticular activating system. Thus the change in alertness reflected by the reduced RT's following a preparatory signal may be attributed in part to a facilitating effect of the ARAS upon the cortex (Lindsley, 1958a, pp. 67–85).

The over-all reduction in mean RT from a nonalerted state (280 msec.) to an alerted state (206 msec.) with a 250–1000 msec. preparatory interval amounts to 74 msec. For similar preparatory intervals the data of Klemmer (1956) and Davis (1940) indicate like reductions in RT. To what may we attribute a reduction of this magnitude? There would seem to be three possibilities: (*a*) receptor and input transmission time, (*b*) output transmission and effector time, (*c*) central integration time. The first two of these are relatively stable and fixed, and subject only to small reductions; the latter appears to be the main source of the reduction in over-all reaction time. Input time, according to Monnier (1952) and Lindsley and Emmons (1958), measured from the incidence of a flash of light upon the eye to the beginning of an evoked potential over the visual cortex, amounts to about 35 msec. Output time, including about 15–20 msec. for transmission from motor cortex to the executant muscle and perhaps 10–20 msec. for muscle shortening, would be no more than 40 msec. over all. The combined input-output time would be about 75 msec. Thus, in a visual RT of 280 msec. at least 200 msec. would have to be attributed to central integration time. Any reduction of the order of 74 msec. would seem likely to involve a reduction in the central integration time rather than the peripheral conduction or transmission time since in the latter there are relatively few synapses to traverse and hence little savings to be effected. Some of the reduction, but probably no more than 10 msec. at most, might be attributed to differences in muscular tension, which exist or develop rapidly following a warning signal as Davis (1940) has shown.

The mechanism which underlies such changes in reaction time seems to involve the reticular formation of the brain stem in at least three ways: (*a*) ARAS facilitation of cortical resolution and transmission of incoming messages (Lindsley, 1958a, p. 70); (*b*) ARAS influence upon the alpha rhythm and its associated excitability cycle (Lansing, 1957; Lindsley, 1958a, pp. 74–78; 1958b; 1958c); and (*c*) descending influences of the reticular formation upon spinal motor outflow (Rossi & Zanchetti, 1957). It appears to us that the first two of these are responsible for the major reductions in RT resulting from preparatory set. The manner in which the first of these influences operates to facilitate cortical transmission is as yet unknown, although there is some evidence that faster response does occur under reticular stimulation (Lindsley, 1958a). The second influence, in which activation or desynchronization of the alpha waves occurs, would seem to affect the waxing and waning excitability cycle associated with alpha wave undulations, tending to eliminate periodic excitability and provide a more continuously and uniformly excitable brain field (Lindsley, 1958a, p. 74). Lansing (1957) found there were fluctuations in RT depending upon the

phase of the alpha wave in which the stimulus fell. Such cyclic variations in excitability could well exist in both sensory and motor cortex and thus increase both the over-all RT and its variability. When such cyclic variations are removed or reduced by ARAS influence through alpha blockade both RT and its variability should be reduced. Indeed this is what has been observed in this study.

Finally the question might be raised as to why the spontaneous disappearance of the alpha rhythm during the nonalerted state is not accompanied by reduced RT's. Although the exact nature of the process responsible for the generation of alpha waves is not known it would appear that local conditions, chemical, electrical, or both, may from time to time change, and lead to spontaneous subsidence of the rhythmic potentials. Such an abolition of alpha waves would, we believe, differ from the situation in which a preceding alerting stimulus, by virtue of its ability to excite the reticular formation, gives rise to persistent activation through the ARAS. The desynchronization of alpha activity through this nonspecific sensory projection system would both block the alpha rhythm and facilitate transmission of impulses in the cortex which have arrived there over specific sensory systems.

Summary

Behavioral alerting or attentiveness is associated with alpha blockade in the electroencephalogram (EEG). Similar blockade and fast, low level activity is produced neurophysiologically by electrical stimulation of the brain stem reticular formation or by sensory stimulation. Psychologically, set or attentiveness is reflected in reduced reaction time. The problem was to relate RT as a psychological measure to EEG activation as a neurophysiological index of alertness or attentiveness.

Visual RT's were recorded in nine normal adult *Ss* during nonalert and alert conditions. Under nonalert conditions visual reaction stimuli were presented during good, poor, and no-alpha states. Under alert conditions a preparatory warning signal preceded the reaction-stimulus by varying foreperiods ranging from 50 to 1000 msec.

Reaction time means and *SD*'s did not differ for the three nonalert conditions, but were markedly reduced when the warning signal produced alpha blocking prior to the visual reaction stimulus. The reduction in RT as a function of the length of the foreperiod interval followed precisely the same time course as the curve showing the degree of alpha blocking as a function of the foreperiod interval. The RT's vary with duration of foreperiod alerting, attaining minimal levels by 300 msec. Since alpha blockade or EEG activation typically occurs within this same interval, both reduced RT and activation are believed to be identified with an alerted or attentive state produced through action of the ascending reticular activating system (ARAS) resulting from sensory stimulation and instructional set.

References

Adrian, E. D., & Matthews, B. H. C. The Berger rhythm: Potential changes from the occipital lobes of man. *Brain*, 1934, **57**, 355–385.

Berger, H. Über das Elektrenkephalogram des Menschen II. *J. Psychol. Neurol., Lpz.*, 1930, **40**, 160–179.

Berger, H. Über das Elektrenkephalogram des Menschen IV. *Arch. Psychiat. Nervenkr.*, 1932, **97**, 6–26.

Davis, R. C. Set and muscular tension. *Indiana Univer. Publ. Sci. Ser.*, 1940, No. 10.

Ellingson, R. J. Brain waves and problems in psychology. *Psychol. Bull.*, 1956, **53**, 1–34.

French, J. D., Hernandez-Peon, R., & Livingston, R. B. Projections from the cortex to cephalic brain stem (reticular formation) in monkey. *J. Neurophysiol.*, 1955, **18**, 74–95.

Klemmer, E. T. Time uncertainty in simple reaction time. *J. exp. Psychol.*, 1956, **51**, 179–184.

Lansing, R. W. Relation of brain and tremor rhythms to visual reaction time. *EEG Clin. Neurophysiol.*, 1957, **9**, 497–504.

Lindsley, D. B. Electroencephalography. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. 2. New York: Ronald, 1944. Pp. 1033–1103.

Lindsley, D. B. Physiological psychology. *Ann. Rev. Psychol.*, 1956, **7**, 323–348.

Lindsley, D. B. Psychophysiology and perception. In *Current trends in the description and analysis of behavior.*

Pittsburgh: Univer. of Pittsburgh Press, 1958. (a)

LINDSLEY, D. B. The reticular activating system and perceptual integration. In D. E. Sheer (Ed.), *Electrical stimulation of the brain*. Austin, Texas: Univer. of Texas Press, 1958. (b)

LINDSLEY, D. B. The reticular system and perceptual discrimination. In *Reticular formation of the brain*. Boston: Little, Brown, 1958. (c)

LINDSLEY, D. B., BOWDEN, J., & MAGOUN, H. W. Effect upon the EEG of acute injury to the brain stem activating system. *EEG Clin. Neurophysiol.*, 1949, **1**, 475–486.

LINDSLEY, D. B., & EMMONS, W. H. Perception time and evoked potentials. *Science*, 1958, **127**, 1061.

LINDSLEY, D. B., SCHREINER, L. H., KNOWLES, W. B., & MAGOUN, H. W. Behavioral and EEG changes following chronic brain stem lesions in the cat. *EEG Clin. Neurophysiol.*, 1950, **2**, 483–498.

MAGOUN, H. W. An ascending reticular activating system in the brain stem. *A. M. A. Arch. Neurol. Psychiat.*, 1952, **67**, 145–154.

MAGOUN, H. W. The ascending reticular system and wakefulness. In *Brain mechanisms and consciousness*. Oxford: Blackwell Sci. Publ., 1954. Pp. 1–20.

MONNIER, M. Retinal, cortical, and motor responses to photic stimulation in man: Retino-cortical time and optomotor integration time. *J. Neurophysiol.*, 1952, **15**, 469–486.

MORUZZI, G., & MAGOUN, H. W. Brain stem reticular formation and activation of the EEG. *EEG Clin. Neurophysiol.*, 1949, **1**, 455–473.

RHEINBERGER, M. B., & JASPER, H. H. The electrical activity of the cerebral cortex in the unanesthetized cat. *Amer. J. Physiol.*, 1937, **119**, 186–196.

ROSSI, G. F., & ZANCHETTI, A. The brain stem reticular formation: Anatomy and physiology. *Arch. ital. Biol.*, 1957, **95**, 199–435.

TEICHNER, W. H. Recent studies of simple reaction time. *Psychol. Bull.*, 1954, **51**, 128–149.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.

(Received July 19, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

PERFORMANCE IN DIFFERENTIAL CONDITIONING AND DISCRIMINATION LEARNING AS A FUNCTION OF HUNGER AND RELATIVE RESPONSE FREQUENCY

K. W. SPENCE, K. P. GOODRICH, AND L. E. ROSS

State University of Iowa

Investigations of the effect of variation of drive level on performance in simple selective (spatial and nonspatial) learning have been highly conflicting. In part the results would appear to be a function of whether or not the number of responses made by *S* to the discriminanda have or have not been equated by means of some forcing technique. Thus three reported studies (Carlton, 1955; Teel, 1952; Teel & Webb, 1951) in which a spatial problem (single T-maze) with equated responses was employed, have found that drive differences do not affect performance in the course of such learning. Similarly, a doctoral investigation by Loess (1952), in which a task with immediate reward on one side and an equal number of delayed rewards on the other was employed, did not find a significant difference between groups deprived of food for 3 and 22 hr. No studies are available in which a nonspatial (e.g., visual) discrimination problem with equating of responses has been used.

Turning to situations in which no attempt has been made to equate the number of responses to the relevant discriminanda, we find that most of the studies have been of the uncontrolled variety in which only free choice trials were used. The results, to say the least, have been somewhat conflicting. Reporting a differential effect of drive in the simple T-maze are the studies of Tolman and Gleitman (1949) and Buchwald and Yamaguchi (1955). In the case of the visual discrimination problem three investigations (Eisman, 1956; Eisman, Asimow, & Maltzman, 1956; Powloski, 1953) have reported a total of four experiments in which there was a significant drive effect. Contrary to these positive findings a brightness discrimination study by Myers (1952) failed to obtain a difference. Likewise, two studies (Carlton, 1955; Champion, 1954) used the free-choice method with the T maze and obtained no evidence of a drive effect. A possible explanation of the negative findings of these last three studies is that the maintenance schedules employed may not have produced an adequate difference in drive level. Tending to support this interpretation was the finding that the running speeds of the subjects to the positive discriminandum failed to reveal significant differences between the high and low drive groups.

Finally, mention should be made of a study by Ramond (1954) in which he gave a differential number of reinforced trials (2:1) to the two responses in a simple spatial situation. It was found that the high drive group performed significantly better than the low drive group. Differences in speed of performance were also obtained in this experiment. Thus the available evidence suggests that when adequate differences in drive obtain, performance will vary with drive differences when there are a differential number of response occurrences to the two discriminanda. On the other hand, if the number of

responses made to the discriminanda are equated by means of some forced trial technique, no difference in performance is obtained.

The present study reports two experiments designed to investigate this possible relation between the effects of drive level on performance in a black-white discrimination learning task and the equation or nonequation of responses to the two discriminanda. The experimental findings will also be related to two alternative hypotheses as to the basis of the development of the habit strength (H) of an instrumental response in learning motivated by an appetitional need.

Experiment I

Method

Subjects.—Forty-four experimentally naive male hooded rats from the colony maintained by the Iowa Psychology Department were used as *Ss*. Ages of *Ss* in the first replication ranged from 132 to 160 days at the beginning of preliminary training; ages in the second replication ranged from 103 to 106 days.

Apparatus.—The black-white discrimination apparatus, the floor plan of which is shown in Fig. 1, was located in a sound-deadened room. The entire apparatus was constructed of ¾-in. white pine, and was covered with clear glass. The entry box and all alleys were 4 in. high and 3½ in. wide. The narrow starting section, designed to remove starting-angle bias from choice responses, was 2 in. wide. The alley section of the apparatus moved on tracks so that any of the four different alley situations could be easily utilized. Two of these presented *S* with a choice situation (Entrances 1 and 2); the other two were straight alleys.

Guillotine doors, which prevented retracing, were located between the entry and starting sections, and at the entrance to all alleys (broken line in Fig. 1). A guillotine door of ¼-in. clear plastic was located at the alley end of the narrow starting section. Raising this door opened a microswitch which started the first of two Hunter silent electronic clocks. The *S*'s interruption of light beam A stopped this clock and started the second; interrupting the second beam B stopped the second clock. The use of Hunter silent relays and clocks eliminated any audible cues arising from the operation of the timing circuit.

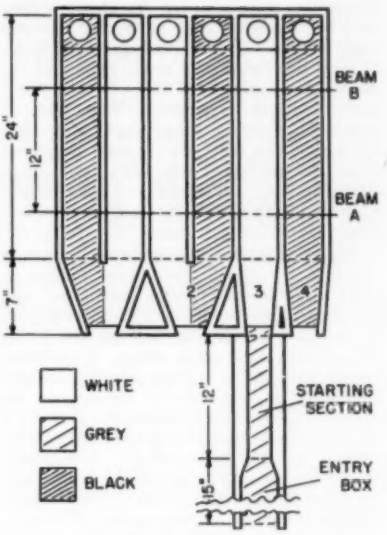

FIG. 1. Ground plan of visual discrimination apparatus.

Alley doors were painted either flat white or flat black as indicated by the shading in Fig. 1 to correspond with the walls and floors of the alleys. All other parts of the apparatus, including the carrying cage placed behind the entry box, were painted flat gray. Alley illumination was provided by a diffused light source 3 ft. above the central axis of the apparatus. The entry box was illuminated with an overhead 7-w. bulb, and a diffuse source of light was located over the carrying cage.

Feeding schedule and preliminary training.—All *Ss* were initially adapted to a 24-hr. limited-time feeding schedule for at least one week prior to preliminary training. During this period *Ss* were allowed 1 hr. of dry food feeding per day in individual cages. The *Ss* were then assigned at random to the two treatment groups, 3 hr. or 40 hr. of food deprivation. Previous work in this laboratory suggested a feeding schedule which would adequately maintain the 40-hr. animals (Davenport, 1956). These *Ss* were fed a wet mash mixture of ground Purina Laboratory Chow, water, and a few drops of cod liver oil, 40 hr. prior to their experimental session. An hour and a half following the beginning of

this session they were fed dry Purina pellets for 1 hr., followed in 5.5 hr. by the next wet mash feeding. This schedule provided 40-hr. *Ss* with two feedings in every 48 hr.; at the same time it permitted the use of a 40-hr. period of food deprivation. The 3-hr. *Ss* were also run every second day. They were fed every 24 hr. for 1 hr., the time at the end of this feeding falling 3 hr. before the experimental session. This feeding consisted of wet mash on running days, dry food on alternate days.

Preliminary training began four days after the *Ss* were shifted from 24-hr. feeding to their experimental feeding schedules, and consisted in placing *Ss* in a 12-in. long gray box which was identical with the last 12 in. of the black and white alleys. Cylindrical food pellets with an average weight of .06 gm. were used as reinforcement throughout the experiment, and were placed in glass food cups set into wooden forms which were painted the color of the alley. In this preliminary training five pellets were available per trial and each *S* received two such trials on each of two consecutive days.

Experimental procedure.—In the experimental phase, all *Ss* received 20 blocks of four trials each, with each block consisting of two runs to white and two to black. One block per day was run on Days 1 and 2, and two per day were run thereafter. The first trial of every block was a choice trial (using alley Entrances 1 or 2). On these trials the white alley occurred half the time on the left and half on the right over the 20 blocks, the order being RLLRRLRRLLRLLRRLRRLL. Following *S*'s choice on Trial 1, he received three straight-alley trials using alley Entrances 3 or 4. The particular sequence of these trials was either "same-different-different" or "different-different-same," where "same" and "different" refer to the brightness chosen on the first trial of the block. Which of these sequences was administered on Block 1 of any daily session was determined by a double-alternation sequence, with half the *Ss* receiving one order and half the other on any one block. If *S*'s Block 2 choice was to the same brightness he had chosen as his Block 1 choice, he received the other straight-alley sequence during Block 2. If his choice was to the other brightness, he received the same straight-alley sequence as during Block 1.

Starting and running times were recorded only on straight-alley trials. Measures, taken on Trials 2 and 4 of every block, were converted to speed measures (reciprocal time in feet per second). The orders discussed above were such as to provide one white and one black measure for each *S* on every block.

On each trial *S* was removed from the carrying cage and placed in the entry box. After *S* had oriented to the clear plastic door of the starting section for 3–4 sec., the door was raised. Reinforcement, consisting of two food pellets, always followed runs to white and never followed runs to black. The *Ss* were removed from the alley as soon as they finished eating on runs to white, and were removed within 10–15 sec. after reaching the end of the black alley. The *Ss* were then returned to the carrying cage along either the left or the right side of the apparatus, this being randomly varied. The next trial began not sooner than 3 min. from the time of return to the carrying cage.

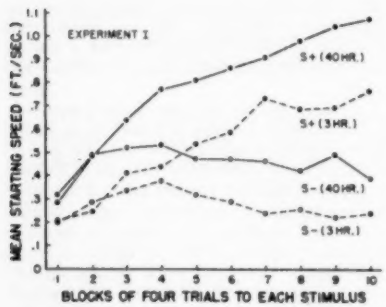

FIG. 2. Curves of response speed to the positive and negative discriminanda during acquisition of discrimination problem under two drive (hunger) levels, with number of responses to the discriminanda equated.

Results

The curves of choice performance for the two drive groups in terms of percentage choice of the positive alley (S_+) on trials that presented both discriminanda remained close together over the 20 blocks of trials, beginning somewhat below 50% and rising to near 100%. Although the 40-hr. *Ss* made slightly more total errors than the 3-hr. *Ss*, the difference was not significant ($t = .41$, $P > .60$).

Starting speed curves for the two drive groups on runs to the two discriminanda are presented in Fig. 2.

Three aspects of these data may be noted. First, the curves for runs to the positive discriminandum (S_+) are negatively accelerated, increasing functions of trials, with the 40-hr. curve being consistently above the 3-hr. curve. Second, the speed curves for the negative discriminandum (S_-) increased slightly at first, then decreased gradually, the high drive group again showing faster speeds than the low drive group. Third, drive appears to have had little or no effect on the difference in response speeds to the positive and negative stimuli, i.e., R_+ minus R_- is the same for groups under high and low drive.

To evaluate statistically the effects of drive upon performance the mean starting speed for blocks of four trials to each stimulus was computed for each *S* and the measures obtained were employed in an analysis of variance involving two within factors (trials and discriminanda) and one between factor (drive). The results of this analysis showed that the main effect of drive level was significant ($F = 29.31$, $df = 1$ and 42, $P < .001$) when tested against the error term based on the variability among *Ss* within drive groups ($MS = 1.48$). The interaction of drive with discriminanda was not statistically reliable ($F = 2.80$, $df = 1$ and 42, $P > .10$) when tested against the error term based on the pooled variability of drive by subject interaction effects ($MS = .13$). These results indicate that drive has an equally facilitating effect on performance to S_+ and S_-, and that the difference in speeds to the two discriminanda, R_+ minus R_-, is not a function of drive level.[1]

[1] The results in both experiments reported here were substantially the same for the running speeds as for the starting speeds. To avoid redundancy, only the starting speed data are presented.

Experiments IIA and IIB

Experiments IIA and IIB differ from each other only in that white was the positive stimulus in Exp. IIA and black was the positive stimulus in Exp. IIB. The description of methodology below pertains specifically to Exp. IIA; it applies also to IIB if "black" and "white" are interchanged throughout. Except as mentioned below, apparatus and procedures for Exp. IIA and IIB are the same as those for Exp. I.

Method

Subjects.—In each of Exp. IIA and IIB the *Ss* were 48 experimentally naive male hooded rats. In Exp. IIA the ages ranged from 151 to 154 days in the first replication, and from 189 to 192 days in the second.

Experimental procedure.—In the experimental phase, all *Ss* received 24 blocks of 3 trials each, each block consisting of two runs to white and one to black. One block per day was run on Days 1 and 2, and two per day were run from then on. The first trial of every block was a choice trial (using alley Entrances 1 or 2). Following *S*'s choice on Trial 1 he received two straight-alley trials using alley Entrances 3 or 4. Following a white choice on Trial 1, *Ss* received the next two trials either in the order white-black or in the order black-white. After a black choice *Ss* always received two trials in the white straight alley. The particular sequence, white-black or black-white, which followed *S*'s first white choice in any experimental session was determined in advance by a double-alternation order. Half of the *Ss* received one of these two sequences and half the other on any one such block. If *S* chose white again in the second block of the experimental session, he received the other straight-alley order during this second block.

Starting and running times were recorded on all straight alley trials, and were converted to speed measures. On blocks in which *S* chose white, one black and one white measure were available. On blocks in which *S* chose black, no speed measures were available.

Results

Performance data on the choice trials in terms of the percentage of

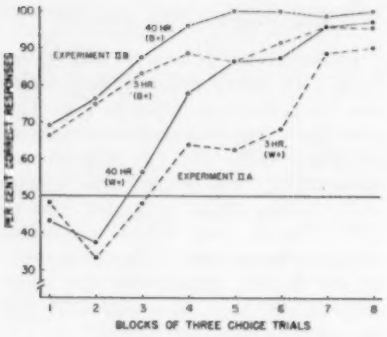

FIG. 3. Acquisition curves on two discrimination problems (W_+ and B_+) under two drive (hunger) levels when responses to the competing discriminanda were not equated.

correct responses for blocks of three trials are presented in Fig. 3. The lower pair of curves (solid circles) represent the data for the *S*s trained positively to the white alley (Exp. IIA), while the upper pair of curves (open circles) are the data for *S*s trained to the black alley (Exp. IIB). It may be seen that in both experiments the high drive group performed at a higher level than the low drive group. In terms of the total number of errors made in the choice trials the drive effect was significant at the .02 level ($t = 2.57$) in Exp. IIA and at the .05 level ($t = 2.12$) in Exp. IIB.

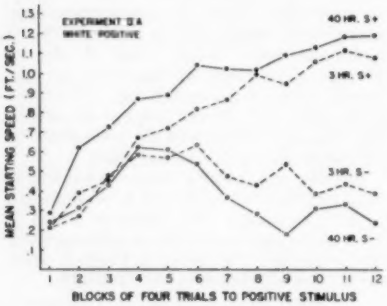

FIG. 4. Curves of response speed to the positive and negative discriminanda under two drive levels with number of responses to two discriminanda not equated (Exp. IIA).

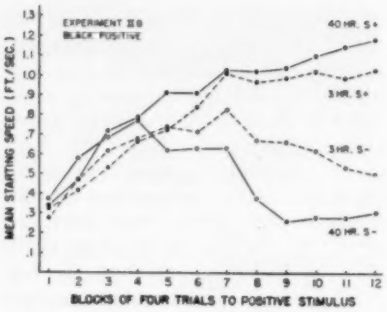

FIG. 5. Curves of response speed to the positive and negative discriminanda under two drive levels with number of responses to two discriminanda not equated (Exp. IIB).

Starting speed curves for the two experiments are presented in Fig. 4 and 5. Several features of these results may be noted. First, although drive again produced a performance difference on runs to S_+, the advantage of the 40-hr. group was not as great as in Exp. I. Second, in contrast to the findings of Exp. I, the speeds to S_- tend to drop more sharply from their maxima. This is especially the case in the 40-hr. groups. Third, the drive differences on runs to S_- reversed from what they were in Exp. I, the 40-hr. *S*s now performing less well than the 3-hr. *S*s over the last part of training. Fourth, in contrast to Exp. I, drive had a clear effect on the differences in speeds to the two discriminanda, i.e., R_+ minus R_- was greater for the 40-hr. group.

It will be recalled that when an *S* in either Exp. IIA or IIB chose the negative discriminandum on any choice trial, no speed measure was available for that discriminandum on that block of trials. Thus during the first half of training there was a considerable amount of missing data for runs to S_-. This has two important consequences.

First, since the missing S_- scores came from *Ss* that chose the negative (incorrect) discriminandum, and which therefore would probably have run relatively rapidly to this stimulus, the performance curves for S_- are probably depressed, especially in the first half of training. Second, these missing data make it impossible to get a criterion measure for some *Ss* on some blocks of trials, precluding a complete analysis of variance by trials as employed in Exp. I. Therefore, it was decided to analyze the data from the last half of training (Blocks 13–24), and to use over-all starting speed means for these 12 blocks as the criterion measures in an analysis of variance involving discriminanda as the within factor and drive as the between factor.

The results of these analyses for Exp. IIA and IIB showed that while the main effect of drive was not significant, the interaction of drive with discriminanda was in both cases ($F = 12.49$, $df = 1$ and 46, $P < .005$; and $F = 19.46$, $df = 1$ and 46, $P < .001$, respectively) when tested against error terms based on the pooled variability of drive by subject interaction effects ($MS = .047, .045$). The statistical evaluation thus reinforces the picture provided by inspection of the performance curves, viz., that the difference in speeds, R_+ minus R_-, is a function of drive level. Examination of Fig. 4 and 5 shows that drive had its greatest effect on performance to the negative discriminandum. The drive differences were tested in both experiments by t tests for the two discriminanda separately. In Exp. IIA the 3-hr. group was significantly superior to the 40-hr. group on runs to S_- ($t = 2.61$, $P < .01$), but for S_+ the drive difference was not reliable ($t = 1.40$, $P > .10$). Likewise, in Exp. IIB the drive effects were significant for S_- ($t = 3.14$, $P < .01$), but not for S_+ ($t = 1.20$, $P > .20$).

Discussion

The findings of these two experiments provide further evidence suggesting that the effect of drive variation on performance in selective learning depends upon whether the number of responses to the two discriminanda are or are not equated. Thus in Exp. I, in which there was an equating of the responses, no difference in choice performance occurred. However, in Exp. II, in which the number of responses to the positive discriminandum was twice that to the negative, a significant difference in choice performance was obtained.

Turning now to the theoretical implications of these experiments, they were initiated primarily to test two alternative assumptions that may be made within the theoretical framework of the senior author (Spence: 1956, 1958). This theoretical schema, originally formulated so as to provide for the derivation and integration of the specific low order laws found to hold in simple conditioning situations, may be tested by deriving new implications concerning phenomena to be expected in more complex behavior situations. The present experiment provided theoretically relevant data for differential instrumental conditioning and selective learning. Thus the speed measures obtained on the forced trials provide data with regard to the course of differential instrumental conditioning, while the choice trials present data about performance during selective (discrimination) learning.

Differential instrumental conditioning.—According to this theoretical formulation, the excitatory potentials of the responses to the two discriminanda (S_+S_-) are given by the following equations:

$$E_+ = H_+ \times (D + K) \qquad [1]$$

$$E_- = H_- \times D - I \qquad [2]$$

where $H = {}_fN_R$ (number of trial runs), $D = {}_fT_d$ (time of food deprivation),

$K = fW_R$ (magnitude of reward), and $I = fn$ (number of nonreinforced trials). The assumption made here, it should be noted, is that habit strength (H) is a function of the number of times a response (locomoting in runway) occurs (N_R). This nonreinforcement (contiguity) assumption concerning the development of H in appetitional learning differs from the reinforcement type of hypothesis, which would assume that H is a function of the number of reinforced responses (N_G).

Now as may be seen from the following equation:

$$E_+ - \bar{E}_- = D(H_+ - H_-) + H_+K + I \quad [3]$$

whether the difference in the strengths of the two excitatory potentials will be a function of drive (D) depends upon whether the two H values are equal or unequal. If they are equal, as is the case when runs are forced equally often to the two discriminanda (Exp. I), then the term $(H_+ - H_-)$ on the right side of Equation 3 becomes zero and the difference between the two E's would be independent of D. That is,

$$E_+ - \bar{E}_- = H_+K + I \quad [4]$$

If, however, these H values are not equal, as is the case in which a greater number of responses occurs to the positive stimulus than to the negative (Exp. II), then the difference $(E_+ - E_-)$ would be a function of D.[3]

Turning now to the implications of the reinforcement hypothesis ($H = fN_G$), since the values of H_+ and H_- in the equated response situation (Exp. I) would not be equal, it follows that the difference between the E's would be expected to be a function of D. It is thus apparent that the implications of the two alternative hypotheses about H differ for the type of equated response situation employed in Exp. I. Their implications with respect to the nonequated response situation (Exp. II), however, do not differ. Since both hypotheses would imply a difference in the habit strengths of the two R's, both would also imply that the difference between their E's will be a function of the level of D present.

As speed of responding (R) is assumed in the theoretical model to be linearly related to the excitatory potential (E) of a response, these theoretical implications concerning the differences between E's can be tested in terms of measurements of the response speeds to the two discriminanda. That is, the speed differences (R_+ minus R_-) are directly proportional to the potential differences (E_+ minus $\bar{E}_-$). Recalling, now, that the speed differences (R_+ minus R_-) in the first experiment did not differ significantly for the two levels of D, it would appear that the contiguity assumption ($H = fN_R$) is favored somewhat over the reinforcement assumption ($H = fN_G$). It should be pointed out, however, that the differences tended to be greater in the 40-hr. group than in the 3-hr. group, but failed of significance ($P > .10$).

As both of the alternative hypotheses predict, the speed differences in Exp. IIA and IIB were significantly greater under the high drive condition. Nevertheless, there is a serious discrepancy between both hypotheses and the speed data of these two experiments. According to the theoretical equation specifying $\bar{E}_-$, the amount of inhibition developed to the negative alley should be the same for both drive groups, for it is assumed that the development of I is independent of T_d. Corroboration of this assumption, it should be noted, was obtained in the first experiment. Thus, as Fig. 2 shows, the decrements in the response speeds to the negative stimulus that develop with nonreinforcement appear to be about the same for the two drive groups. Quite in contrast, however, are the data of Exp. IIA and IIB. As Fig. 3 and 4 reveal, the response speeds to the negative alley decreased at a faster rate in the case of the high drive group in both

[3] According to the contiguity assumption, however, this difference between the habit strengths of the two responses after first increasing would begin to decrease and eventually become zero. See discussion in Spence (1956, pp. 208–214).

experiments. These latter data suggest the hypothesis that the development of I is a function of drive level and, hence, that the equation for $\bar{E}_-$ should be modified as shown in the following:

$$\bar{E}_- = D(H - I) \qquad [5]$$

This latter assumption, however, will not fit the findings of Exp. I, for the results of the two studies are in fundamental disagreement so far as the effects of drive differences on the strength of nonreinforced responses are concerned. It is perhaps obvious that we need to obtain much more knowledge than we now possess concerning the variables affecting the development of response decrement with nonreinforcement. Unfortunately, this problem has been badly neglected in conditioning experiments with the consequence that such an empirically based theory as the present one is weakest in this area.

Selective learning.—The theoretical relation between choice behavior, percentage of choice of the positive (reinforced) response, and the magnitude of the difference between the competing E's (E_+ minus $\bar{E}_-$) is more complicated than in the case of response speeds. As Spence (1956, 1958) has shown, it is possible by means of a composition rule (or law) concerning the interaction of the competing E's and the laws or postulates concerning the variables in the individual response systems, particularly those relating to oscillatory inhibition (I_o), to derive the implication that the percentage occurrence of the response (R_p) to the positive stimulus is a normal integral function of the magnitude of the difference between the strengths of the competing excitatory potentials.

Unfortunately, as has been pointed out (Spence: 1956, 1958) R_p is not only a function of the size of the difference between the competing E's, but depends also on their absolute levels above the threshold. Without going into detail, it may be shown that the contiguity hypothesis would predict in the equated response situation (Exp. I) that at no level of training would the high drive group be expected to perform at a higher level of correct choice than the low drive group (Spence, 1958). In Exp. II, on the other hand, this particular hypothesis would predict a difference in choice behavior for the two drive levels. Beyond the earliest stage of training, at least, the performance of the high drive group would be expected to be above that of the low group. As we have seen, the experimental findings have been in accord with these theoretical predictions.

The implications of the reinforcement hypothesis for the equated response situation (Exp. I) present a somewhat more complex picture (Spence, 1958). According to the theoretical model employing this assumption, the effects of drive differences will vary with the absolute level of the drives being compared and the level of the habit strengths of the two responses (stage of training). At extremely low levels of drive and training the relatively weaker drive group would be expected to perform at a higher level than the stronger drive group. At the upper range of drive levels and at relatively high levels of H, the performance of the higher drive group should be above that of the relatively weaker drive group. At intermediate values of D and H it may be predicted that there would be little or no difference in the performance of two groups differing in drive level. Since we are unable to specify at what levels of these parameters our experiments fall, it is impossible to make a specific prediction. About all that can be said with regard to the reinforcement hypothesis is that one would expect the drive difference to have essentially the same effects in Exp. I and II. The fact that this finding did not obtain is not favorable to this assumption.

It is apparent from this discussion that a much more extensive program of research involving a wide range of drive levels is needed before one would be in a position to make a decision between these two alternative assumptions. The present data, insofar as they provide further evidence for the conclusion that the effect of drive variation in this type of learning situation depends upon whether the

competing responses are or are not equated, tend to support the contiguity more than the reinforcement hypothesis. The former hypothesis would predict such a relation, whereas the latter would not.

Summary

The present study reports two experiments designed to investigate the effects of varying drive level on differential conditioning and selective learning.

In Exp. I, two groups of 22 hooded rats each (3-hr. and 40-hr. of food deprivation) were run for 20 blocks of trials in a black-white discrimination apparatus. The one choice and three forced trials in each block were arranged so that *S* ran twice to white (always reinforced) and twice to black (never reinforced). Choice trials provided data on selective learning; forced trials provided speed data relevant to differential instrumental conditioning. Curves of percentage choice showed no difference between drive groups, and speed curves showed that the difference between the speeds to the two discriminanda was not a function of drive level.

In each of Exp. IIA and IIB, two drive groups of 24 *S*s were run for 24 blocks of trials in the discrimination apparatus. Each block consisted of one choice trial and two forced trials such that there were two runs to the positive discriminandum and one to the negative. White was the positive stimulus in Exp. IIA, black in Exp. IIB. In both experiments, curves of percentage choice showed a significant advantage in favor of the 40-hr. group, and the speed curves showed that the difference between the speeds to the two discriminanda was significantly greater for the 40-hr. *S*s.

These findings provide further evidence for the generalization that performance in selective learning is independent of drive level when responses to the two discriminanda are equated, but varies with drive when the responses are not equated. The results were examined in relation to two alternative hypotheses of habit growth, in the context of an extension of Spence's conditioning theory to selective learning.

References

Buchwald, A. M., & Yamaguchi, H. G. The effect of change in drive level on habit reversal. *J. exp. Psychol.*, 1955, **50**, 265–268.

Carlton, P. L. The effect of time of food deprivation on selective learning. Unpublished Ph.D. dissertation, State Univer. of Iowa, 1955.

Champion, R. A. Drive-strength and competing response tendencies. Unpublished M.A. thesis, State Univer. of Iowa, 1954.

Davenport, J. W. Choice behavior as a function of drive strength and rate of learning. Unpublished Ph.D. dissertation, State Univer. of Iowa, 1956.

Eisman, E. An investigation of the parameters defining drive (D). *J. exp. Psychol.*, 1956, **52**, 85–89.

Eisman, E., Asimow, A., & Maltzman, I. Habit strength as a function of drive in a brightness discrimination problem. *J. exp. Psychol.*, 1956, **52**, 58–64.

Loess, H. B. The effect of variation of motivational level and changes in motivational level on performance in learning. Unpublished Ph.D. dissertation, State Univer. of Iowa, 1952.

Myers, T. I. An experimental investigation of the effect of hunger drive upon the brightness discrimination learning of the rat. Unpublished Ph.D. dissertation, State Univer. of Iowa, 1952.

Powloski, R. F. The effects of combining hunger and thirst motives in a discrimination habit. *J. comp. physiol. Psychol.*, 1953, **46**, 434–437.

Ramond, C. K. Performance in selective learning as a function of hunger. *J. exp. Psychol.*, 1954, **48**, 265–270.

Spence, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.

Spence, K. W. Behavior theory and selective learning. In M. R. Jones, (Ed.), *Nebraska symposium on motivation*. Lincoln: Univer. of Nebraska Press, 1958.

Teel, K. S. Habit strength as a function of motivation during learning. *J. comp. physiol. Psychol.*, 1952, **45**, 188–191.

Teel, K. S., & Webb, W. B. Response evocation on satiated trials in the T-maze. *J. exp. Psychol.*, 1951, **41**, 148–152.

Tolman, E. C., & Gleitman, H. Studies in spatial learning: VII. Place and response learning under different degrees of motivation. *J. exp. Psychol.*, 1949, **39**, 653–659.

(Received July 19, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

ON THE PREDICTION OF OCCURRENCE OF PARTICULAR VERBAL INTRUSIONS IN IMMEDIATE RECALL[1]

JAMES DEESE

Johns Hopkins University

In the past few years the author has published a series of experiments dealing with the frequency and order of emission of particular items in recall. These experiments demonstrate that both the order of emission of items in free recall and the occurrence of intrusions from other experimental lists may be determined by the context of a given list. If a list is learned by serial anticipation, most intrusions in serial recall, even after interpolation, come from other positions within the same list (Deese & Hardman, 1954; Deese & Marder, 1957). If, however, a list is organized into paired-associates, intrusions after interpolation are more likely to come from other lists, particularly if the same stimulus items are used in all lists (Deese & Hardman, 1954). In free recall, the order of emission of items within a list depends upon the strength of the serial organization in the list (Deese, 1957; Deese & Kaufman, 1957). Lists which have no serial or sequential organization are recalled in order of the associative strength of the items. Lists in which sequential organization is induced by statistical dependencies between items are recalled in serial order. Thus, sequentially organized lists in free recall behave much as do lists learned by serial anticipation.

Despite these and similar generalizations it has been impossible to describe completely the organization of lists in recall because of the "unpredictable" occurrence of extra-list intrusions. These are items which occur during recall but which come from none of the lists presented to *S*s. A plausible ex post facto account can be given for most extra-list intrusions after they have occurred, but heretofore no scheme has been developed for predicting their occurrence. The purpose of the present paper is to describe a technique for predicting the occurrence of such extra-list intrusions and to demonstrate that they are accounted for in terms of simple association, or, more specifically, the associative context of the lists presented to *S* for recall.

There are two problems inherent in the study of extra-list intrusions. These are (*a*) producing lists which reliably result in particular extra-list intrusions, and (*b*) finding variables which predict the relative frequency of occurrence of those intrusions. In the present experiment both of these problems have been solved by the use of the word-association technique, and, therefore, the present analysis demands the assumption that word-association norms may be used as indices of the strength of verbal stimulus-response associations in the population from which *S*s have been drawn. This assumption is justified to the extent that word-association norms are able to predict the occurrence of particular extra-list intrusions.

METHOD

The experiment consisted of two parts. In one, the data on the occurrence of particular extra-list intrusions were gathered. In

[1] The research reported in this paper was in part supported by funds from the National Science Foundation Grant NSF-61369.

the other, word-association frequencies for the items used in the recall tests were obtained. Here the items from the recall lists served as stimulus items in a free association test.

In the study of immediate recall, 50 undergraduate students at Johns Hopkins University were used as *Ss*. All *Ss* were tested individually. Each *S* was tested on 36 lists of 12 words each. The *Ss* were instructed to listen carefully to each list and to recall orally the items from each list at the conclusion of the reading of the list. The lists and instructions for immediate free recall were presented by tape recording. The oral recalls from *Ss* were recorded by tape and transcribed later. Five different orders of lists were used; these were obtained by independent random assignment of the 36 lists to positions in testing. The assignment was made on the basis of tabled random numbers. All five orders were recorded at the outset of the experiment, and these were presented to *Ss* in rotated order as *Ss* were chosen from the available pool.

Each list consisted of the 12 highest frequency responses given to a particular stimulus on the Kent-Rosanoff word-association lists. The frequencies were obtained from the Minnesota norms for the Kent-Rosanoff items (Russell & Jenkins, 1954). The particular 36 lists are taken from Items 1, 2, 3, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 33, 35, 36, 37, 38, 39, 40, 41, and 43 of the Minnesota norms. These particular 36 lists were chosen on the basis of preliminary observations which indicated that some of the lists very reliably produced the tendency for *Ss* to recall the stimulus word associated with the list (a word they had not heard in the list) and others of the lists never or very infrequently produced a tendency to recall the stimulus word. Thus a considerable range of tendency to emit an intrusion in recall resulted.

In order to predict the occurrence of these particular intrusions in recall it was necessary to obtain word-association norms with the words on the recall lists used as stimuli. Therefore, word association tests were administered to samples of *Ss* drawn from the same population as that used in the recall tests but consisting of different individuals. Since there were many items to be used in the word-association measurement, the items were divided into two groups and administered to two separate samples of 50 *Ss* each. There were 200 items on each test, including some filler items. These tests were administered in group form with the following printed instructions on the cover of the test booklet (these instructions were also read by *E*): "This is a free association test. On the next four pages you will find a list of words. Look at each word and write down in the column next to it a word that it makes you think of. Write only one association to each word. Sometimes you will think of a word that you have already given as a response to another stimulus word. Always write down this word; in brief, do not let your responses to earlier items influence how you respond to later items. Use only a single word. Do not skip any word."

The major data obtained from these word-association tests consisted of percentage frequency of occurrence of the Kent-Rosanoff stimulus word appropriate to a particular list as an associative response to words from that list. For example, "Chair" is the highest frequency response to "Table" on the Minnesota norms of the Kent-Rosanoff. In the present case the frequency with which "Table" occurs as a response to "Chair" was tabulated. For each list used in recall, the mean percentage frequency occurrence of the Kent-Rosanoff word as a response to all the items on the list was calculated for each list. These mean percentage frequencies constitute the independent variable in the present study.

Since the obtained association frequencies in the present study are based upon responses from only 50 *Ss*, a small number for word-association studies, it is fortunate that some indication of the reliability of these associations may be obtained by comparing them with the frequencies reported in the Minnesota norms for those stimulus items common to both the present lists and the Kent-Rosanoff. The Minnesota norms are based upon a sample of 1008, obtained from a population similar to that from which *Ss* for the present study have been drawn, so the obtained frequencies from the present study should correlate with those obtained from the Minnesota norms. There were enough common stimulus items between the lists such that there were 102 comparable associations which involved the critical items (responses which were the stimulus words for the 36 Kent-Rosanoff items). A number of these were zero frequency for both the Kent-Rosanoff and the present tables (e.g., while "man" occurs as a response to "short," "short" never occurs as a response to "man"). These common zero association frequencies were discarded in the comparison, however, on the grounds that they could

unduly inflate the obtained correlation. Consequently, a product-moment correlation was computed between the obtained percentage frequency of association for 88 different critical responses to stimuli common to both sets of data. The resulting r was .778. The mean percentage frequency for these items on the Kent-Rosanoff norms was 16.0, and the mean percentage frequency from the present data was 15.4. Therefore, despite the comparatively small sample upon which the norms used in the present study are based, they seem to correspond well to those obtained from a much larger sample from a similar population. Since the average association score for each list used in the recall study is the mean percentage frequency with which the critical word occurs to 12 different stimulus items, even greater stability is achieved in the independent variable as the result of combination.

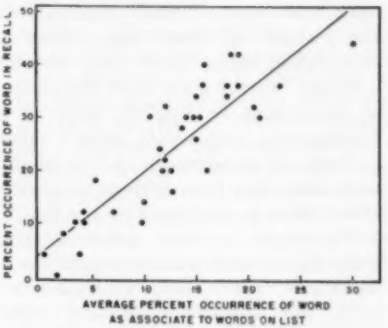

FIG 1. Percentage occurrence of a particular word as an intrusion in immediate free recall as a function of the mean association strength of that word to the words on the list.

RESULTS

Table 1 shows the percentage frequency with which the Kent-Rosanoff stimulus word occurs as an intrusion in immediate free recall of the list composed of response items to that stimulus word. These vary from 0% for the stimulus word "Butterfly" to 44% for the stimulus word "Sleep." The problem in the present study is to determine the extent to which the tendency to emit these particular words as intrusions in recall is predictable from the frequency with which these words are given as associations to words on the list. Figure 1 shows the relationship between the percentage frequency occurrence of the stimulus word as an intrusion in recall and the mean percentage frequency of the stimulus word as an association to items on the list. As can be seen, there is an apparent regression between these variables. The product-moment correlation for these data is .873. With 35 *df*, it is significant well beyond the .1% level. If the obtained correlation is used as the best estimate of the true correlation, it follows that 76% of the variance in probability of intrusion is determined by variation in mean association strength.

TABLE 1

PERCENTAGE OCCURRENCE OF STIMULUS WORD AS INTRUSION IN IMMEDIATE RECALL (N = 50)

Stim. Word	% Intrusion	Stim. Word	% Intrusion
Table	18	Whistle	4
Dark	10	Cold	34
Music	30	Slow	30
Man	32	Wish	12
Deep	14	River	20
Soft	40	White	16
Mountain	36	Beautiful	8
House	12	Window	30
Black	28	Rough	42
Mutton	4	Foot	36
Hand	22	Needle	42
Short	30	Red	24
Fruit	20	Sleep	44
Butterfly	0	Anger	34
Smooth	26	Carpet	10
Command	10	Girl	32
Chair	36	High	30
Sweet	36	Sour	20

Since the regression is apparently linear, a least-squares fit of the relationship was made. This yielded $Y = 3.7 + 1.63X$. The slightly positive origin suggests that intrusions

may have some "free" strength in the context of these lists. This is reasonable, since they are, on the average, words with high frequency of occurrence in English, and, as a consequence, one would expect them to turn up occasionally as intrusions, even when they have little or no association value to the items on the list.

The slope constant predicts that if the mean association strength for a particular word were increased to 60%, the probability of that word occurring as an intrusion would be 1.0. The author has tried to construct lists which would yield a particular word as an intrusion nearly always, but it does not seem possible to do so with the association norms currently available and if the lists are to be 12 items in length. It is very possible, however, that if association norms were available for a very large number of words, such a list could be constructed. If so, it would provide a definitive test of the linearity of regression of probability of intrusion upon association value.

It is possible to argue that the relationship described above is actually the result of other, uncontrolled variables. The most important of these are (*a*) the frequency of usage of the words predicted as intrusions and (*b*) the frequency with which these words occur in other prior lists. Both of these variables were examined as possible predictors of frequency of intrusion.

The L count from the Thorndike and Lorge (1944) word count was used as a measure of frequency of usage. The correlation between L-count frequencies and percent intrusions was .059. It might be argued that the correlation would have been higher if the L-count frequencies had been transformed logarithmically; however, both the highest and lowest frequency intrusions were among the very lowest L-count words. The low correlation is the result of scatter rather than curvilinearity in the plot. It is possible, however, that there would have been a genuine though small effect of frequency of usage if an adequate sample of words of different frequency of usage had been employed. In the present experiment only three of the predicted intrusions had L frequencies below 100, and none had L frequencies below 22. Therefore, while frequency of usage cannot account for the frequency with which intrusions occur, it is very possible that an investigation of frequency of usage itself would reveal a correlation between it and intrusion in recall. The investigation of this possibility awaits the development of word association norms which sample words adequately from a wide range of frequency of usage.

Some of the intrusions occur as actual items in other lists. This is true for 23 of the 36 predicted intrusions. Therefore, it might be argued that such intrusions were the result of a proactive effect. On general grounds this is unlikely, since the technique of a single presentation and a single immediate recall trial minimizes influences from list to list. However, the frequency of occurrence of each intrusion as an actual item in all prior lists was tabulated for each *S*, and the summed frequencies for all *S*s were correlated with the frequency of intrusion. The resulting r was .162, which is not statistically significant. It is possible that proactive effects do exist, but they have not been satisfactorily demonstrated here, and they do not account for the obtained correlation between association frequency and frequency of intrusion.

In addition to the actually predicted intrusions in particular lists, the same words occasionally appear as intrusions in other lists, and, additional, unpredicted words also appear. These appear, however, very infrequently. The predicted intrusions occur, for all lists, with an average frequency of 24.2%, while the unpredicted intrusions occur with an average frequency of only 3.2%. Of these unpredicted intrusions, the word "girl" (a predicted intrusion for one list) occurs three times, each occurrence on a different list, and the word "music" (also a predicted intrusion for one list) occurs three times within one list, the list for which the intrusion "whistle" is appropriate. No other unpredicted word occurs as an intrusion more than twice, although for nearly every list at least one unpredicted intrusion occurs. The average frequency of 3.2% for these unpredicted intrusions agrees well with the value for the origin of the fitted curve in Fig. 1, if the possibility of occasional proactively determined intrusions and occasional unpredicted associations is taken into account.

Discussion

The results of the present experiment indicate that the probability of a particular word occurring as an intrusion in immediate free recall of a list of words may be predicted from the tendency for the intruding word to occur as a response in free association to the items on the list. It is implicit that such association frequencies represented previously learned habits. Therefore, the present study has some implications for the influence of previous habits upon retention, an influence the importance of which has recently been pointed out by Underwood (1957).

The question immediately arises, may the results of this study be generalized to the occurrence of intrusions in recall at large? With the exercise of some caution, some generalizations may be made which, in turn, suggest some interesting hypotheses. The particular relationships reported in this paper are, of course, limited by the conditions of the experiment. The important point, however, is that the effect occurs, and, furthermore, for the data reported here it appears to be a linear effect. This implies that there is no interaction among the several words. The probability of intrusion is very well predicted by the simple mean, irrespective of distribution, of the associative tendencies for the intrusion to be elicited by the test words. Because of the very high proportion of variance predicted by the mean associative strength, relatively little variance could be predicted from any possible interactions. Therefore, the present data seem to warrant the hypothesis that the probability of occurrence of an intrusion in recall is proportional to the average association strength of that item in the context of the material being recalled and is relatively independent of the distribution of association frequencies among the various items on the list.

If this hypothesis has general validity, it provides an important mechanism for analysis and description of the kinds of elaboration in memory described by Bartlett (1932). It implies that in the process of recollection, words and concepts associated with remembered items will be added. Likewise, all other things being equal (ignoring, for example, serial position effects), items which are forgotten will be items which have low probabilities of being elicited by remembered items from the list. Thus, it may well be possible that some of the effects described as leveling and sharpening may turn out to be simply a matter of the strength of contextual associations to the items to be remembered.

With this in mind it is possible to interpret "clinically" some of the material presented by Bartlett (which, after all, is exactly what Bartlett did). For example, in the celebrated "War of the Ghosts" the word "war" appears four

times near the beginning of the story, but the words "battle" and "enemy" never appear. In one of the initial reproductions presented by Bartlett the word "war" appears, but also the words "battle" and "enemy" appear twice. In general, if the word "war" appears in recall, the words "battle" and "enemy" are added. These latter words are both high frequency associations to "war." Such an analysis proves nothing, of course, but it does suggest the possibility of constructing meaningful passages containing repeated words with known probabilities of eliciting other words as associations and then examining the reproductions for the frequency of occurrence of these associations as intrusions.

Even more intriguing is the possibility that some of the well-known experiments, the results of which are usually described as the effects of attitude and motive upon retention, may be predictable from the association frequencies to critical words yielded by the target populations. For example, one of the best known studies (Levine & Murphy, 1943) compared the retention of pro- and anti-communistic material in groups of *Ss* known to be either pro- or anti-communistic. The results of the present study suggest that at least some of the differences in kinds of material retained by such groups would be predicted from differences in association frequencies for critical words between the groups.

Thus, the hypotheses arising out of the empirical relationship found in this study suggest ways of interpreting a wide variety of problems of patterning and organization in memory in terms of elementary associative frequencies. At the very least, the results of this study suggest that experiments in which different groups of *Ss* are compared for their ability to remember specific items should either present data on association frequencies for the groups concerned or make certain that such association frequencies are well controlled.

Summary

Lists consisting of 12 words each were presented to 50 *Ss* for a test of immediate recall. In the recall of these lists, particular words occurred as intrusions which varied in frequency from 0% for one list to 44% for another. Data gathered on word-association frequencies clearly showed that the probability of a particular word occurring in recall as an intrusion was determined by the average frequency with which that word occurs as an association to words on the list. The correlation between probability of intrusion and mean association value was .873. The regression, over the range examined, was linear, and this suggests the hypothesis that the probability of occurrence of a particular word as an intrusion in recall is proportional to the average association strength of that word to the words on the list.

References

Bartlett, F. C. *Remembering.* New York: Cambridge Univer. Press, 1932.

Deese, J. Serial organization in the recall of disconnected items. *Psychol. Rep.,* 1957, **3,** 577–582.

Deese, J., & Hardman, G. W., Jr. An analysis of errors in retroactive inhibition of rote verbal learning. *Amer. J. Psychol.,* 1954, **67,** 299–307.

Deese, J., & Kaufman, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. *J. exp. Psychol.,* 1957, **54,** 180–187.

Deese, J., & Marder, V. J. The pattern of errors in delayed recall of serial learning after interpolation. *Amer. J. Psychol.,* 1957, **70,** 594–599.

Levine, J. M., & Murphy, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.,* 1943, **38,** 507–517.

Russell, W. A., & Jenkins, J. J. *The complete Minnesota Norms for responses to 100 words from the Kent-Rosanoff Association Test.* Tech. Rep. No. 11, Contract N8ONR-66216 between Of. of Naval Res. and Univer. of Minnesota.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Bur. of Publ., Teacher's Coll., Columbia Univer., 1944.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.,* 1957, **64,** 49–60.

(Received July 7, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

ABSOLUTE JUDGMENTS OF ODOR INTENSITY[1]

TRYGG ENGEN AND CARL PFAFFMANN

Brown University

This paper is concerned with the ability of the human *S* to transmit information about odor intensity. The problem is to determine the number of stimuli, in this case, the concentration dilution values of an odorant, that *S* can identify with a minimum of error when they are presented singly according to the method of absolute judgment. Miller (1956) has compared such data for auditory, visual, gustatory, and tactual stimuli, but no results have been reported for olfaction.

The present experiments were concerned with the following questions: (*a*) What is the greatest number of intensities of an odor that can be identified correctly by rank order? (*b*) How is the accuracy of these identifications influenced by the selection of odorants, the size step between stimuli, intensity of stimuli, practice, and the number of stimulus categories presented to the observer?

METHOD

The odor series.—The odor stimuli were prepared according to a procedure suggested by Beck, Kruger, and Calabresi (1954). Mallinkrodt's Benzyl Benzoate (USP XIII, 387 control TSN-1) was used as a diluent. The odorants used in the experiments were amyl acetate, n-heptanal, n-heptane, and phenylethyl alcohol. The diluent and the odorant were mixed in a 10 × 75 mm. test tube with a pipette. A geometric dilution series was prepared by progressively halving the dilution value of the odorant such that Stimulus 1 of the series contained 100% of the odorant, Stimulus 2 contained 50% of the odorant and 50% of the diluent, Stimulus 3 contained 25% of the odorant and 75% of the diluent, etc. For all four odorants the threshold stimulus contained .00003% or less of the odorant, i.e., Stimulus 16 or higher. The stimuli used in the experiment were selected from this series and were renewed frequently. Each test tube, containing 1 cc. of odorant plus diluent, was stopped with an aluminum wrapped cork which was removed only for sniffing. Other studies have shown that sniffing is a reliable method for both scaling and threshold measurements in olfaction (Beck et al., 1954; Garner & Hake, 1953).

Subjects.—Judgments were obtained from eight *S*s, three men and five women. They were, with one exception (a woman approximately 40 yr. of age), college juniors and were paid at the usual student rate of $1.00 per hour. All had previous experience in rating and ranking odor intensities.

Procedure.—The experiments were performed in a well-ventilated and air-conditioned room maintained at a temperature of about 24° C. In the first part of the experiment *S* learned to identify stimulus intensities. A rack with five stimuli rank-ordered from strong to weak were presented to *S*s, and they were instructed to learn the correct rank order of each stimulus so that they could identify it later when it was presented singly. The *S* would then sniff the content of each test tube, compare it with the others until he was satisfied no improvement would result from further practice. The stimulus rack was then returned to *E* who placed it out of sight behind a screen. This procedure was repeated at the beginning of each day.

Each of the five stimuli was then presented singly in a random sequence, and *S* was given unlimited time to identify its rank order of intensity among the five stimuli. Each stimulus was presented 50 times for a total of 250 judgments for each set of five stimuli. The correct rank order was indicated by *E* immediately after *S* made his judgment. After a period of about 12 min., *S* was allowed to rest while one or two other

[1] This research was supported in part by a grant from the General Foods Corporation. This report presents part of a paper read at the 22nd Annual Meeting of the Industrial Hygiene Foundation of the Mellon Institute, October 30, 1957. The present report was also presented at the meeting of the Eastern Psychological Association in Philadelphia, April 12, 1958.

TABLE 1

AVERAGE AMOUNT OF INFORMATION IN BITS OBTAINED FOR FIVE ALTERNATIVE STIMULI WITH 100, 50, 25, 12.5, AND 6.25% DILUTION CONCENTRATION VALUES, FOR EACH OF FOUR ODORANTS

Odorant	*N*	Mean	*SD*
Amyl acetate	5	1.52	.37
n-Heptanal	4[a]	1.53	.16
n-Heptane	5	1.51	.33
Phenylethyl alcohol	5	1.58	.27

[a] One woman became nauseated from this odorant and testing had to be discontinued.

Ss were run in a similar manner. Each *S* thus worked and rested for a maximum of 3 hr. per day, 3 to 4 days per week. Approximately one week was required to complete the judgments on one set of five stimuli. An effort was made to schedule *Ss* at the same time of the day, and they were asked not to use perfume on experimental days nor to eat or smoke for 1 hr. before an experimental session.

Analysis of data.—The absolute judgments lend themselves to the informational analysis described by Garner and Hake (1951). These writers have shown that the analysis of absolute judgments in bits per stimulus yields an index of the largest number of stimuli that can be judged without error.

RESULTS AND DISCUSSION

Table 1 presents the results from the first experiment on absolute judgments of amyl acetate, n-heptanal, n-heptane, and phenylethyl alcohol in dilution series of 100, 50, 25, 12.5, and 6.25%. The four sets of stimuli were presented at random to each of five *Ss*. The information transmitted for each odorant was computed individually for each *S*. The *SD*'s of the means give a rough indication of the relatively large individual differences.

The results indicate that with a brief period of practice *S* is able to transmit about 1.5 bits of information about intensity differences in odor. These results are in close agreement with those obtained for taste with saline and sucrose solution differences (Beebe-Center, Rogers & O'Connell, 1955) but somewhat lower than those obtained for auditory and visual stimuli varying in only one dimension (Eriksen, 1954; Eriksen & Hake, 1955; Miller, 1956). This suggests that differences in intensities of taste and odor appear to be somewhat less distinctive than are those in the other sense modalities. Thus, whereas in taste and odor an *S* can learn to discriminate and remember three categories (the antilog of 1.53 bits), the number of categories in audition, for example, is approximately five.

Table 1 indicates that essentially the same results were obtained with four different odorants. It is not possible, of course, to conclude that a reliable difference might not be found with other odorants. However, preliminary work with several additional odorants indicate that such differences would be very small.

The concentration dilution value of the stimuli and the size step between the stimuli do have a small but reliable effect on such judgments. Twelve sets of stimuli were prepared by using three different dilution series

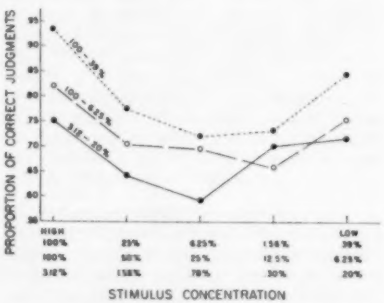

FIG. 1. Proportion of correct absolute judgments of odor intensity as a function of the position of the stimulus in the series, size step, and concentration dilution value.

of the four odorants described above, i.e., series of (*a*) 100, 50, 25, 12.5, and 6.25%, (*b*) 3.12, 1.56, .78, .39, and .20%, and (*c*) 100, 25, 6.25, 1.56, and .39%. The 12 sets were assigned in a random sequence according to odorant, step size, and concentration dilution value. Approximately nine consecutive weeks were required to complete judgments on all 12 sets of stimuli with each *S*.

Average results for the four odorants combined were 1.6, 1.4, and 1.7 bits for Dilution Series *a*, *b*, and *c*, respectively. The results are presented in Fig. 1 as the proportion of correct judgments of the rank order of each stimulus dilution concentration. First, relatively weak stimuli (low concentration values) are less likely to be judged correctly than are relatively strong stimuli. That is, the two lower curves in Fig. 1 show that the stimulus set consisting of 3.12, 1.56, .78, .39, and .20% solutions are judged less accurately than is the stimulus set consisting of 100, 50, 25, 12.5, and 6.25% dilutions. In both sets the step between adjacent stimuli was 50% dilution. Secondly,

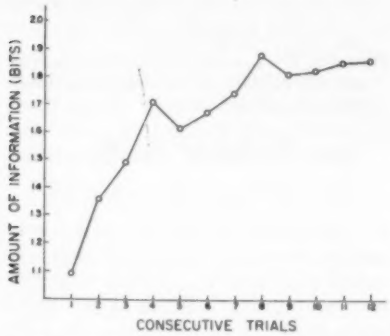

FIG. 2. Practice effects on absolute judgments of odor intensity measured by amount of information transmitted in 12 consecutive trials. Each trial is on a different set of five stimuli.

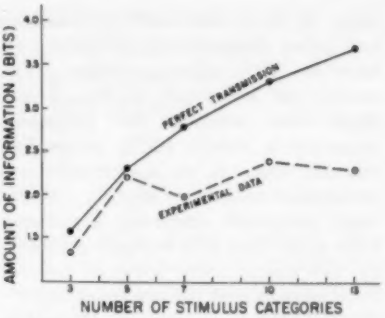

FIG. 3. Amount of information transmitted about odor intensity by a well-practiced observer as a function of the number of stimulus categories presented for judgment.

the stimulus set consisting of the 100, 25, 6.25, 1.56, and .39% solutions, i.e., with a larger dilution step between the stimuli, yielded better results than did the other two sets. It is interesting to note that the usual end effect appears in the judgment of odor as in other sense modalities.

The first experiments showed a practice effect, i.e., improvement in accuracy on successive trials with the several sets of stimuli. In order to obtain a measure of the practice effect the 12 sets of stimuli were presented as above to three new *S*s. The results presented in Fig. 2 show that *S*'s performance, measured in bits, improved appreciably up to Trial 8, where it levels off at about 1.9 bits. This indicates that with practice *S* can discriminate about four stimulus categories with a minimum of error.

These practice results are based on group averages. The best of these *S*s reached nearly perfect performance with five stimuli. To assess "maximum channel capacity" (Alluisi, 1957) and the degree to which such performance is influenced by the number of categories available for judgment, this well-practiced *S* was presented

with 3, 5, 7, 10, and 13 stimulus categories consecutively of amyl acetate in 50% dilution steps. The results are presented in Fig. 3 and show that whereas the judgment accuracy is nearly 100% up to five stimuli, there is no appreciable improvement with more stimuli. Moreover, judgment efficiency is reduced with more than five stimuli, as shown by the discrepancy between perfect transmission and the obtained experimental data. It seems reasonable to assume that the above results were not determined by the number of stimulus categories used in the experiments. This is in agreement with the general findings in other sense modalities (Ericksen & Hake, 1955; Miller, 1956) that as the number of alternative stimuli and responses is progressively increased the amount of information transmitted will reach an asymptote.

Summary

Results obtained with the method of absolute judgment suggest that a relatively unpracticed *S* can identify correctly three levels of intensity of an odorant and that a well-practiced *S* can identify about four levels. There is only a small though reliable improvement in the accuracy of these identifications with increase in the intensity level of the stimulus series and the size of the step between the stimuli. Increasing the number of alternative stimuli presented beyond five did not influence the results appreciably. Essentially the same results were obtained with four odorants, amyl acetate, n-heptanal, n-heptane, and phenylethyl alcohol.

References

Alluisi, E. A. Conditions affecting the amount of information in absolute judgments. *Psychol. Rev.*, 1957, **64**, 97–103

Beck, L. H., Kruger, L., & Calabresi, P. Observations on olfactory intensity: I. Training procedure, methods and data for two aliphatic homologous series. *Ann. N. Y. Acad. Sci.*, 1954, **58**, 225–238.

Beebe-Center, J. G., Rogers, M. S., & O'Connell, D. N. Transmission of information about sucrose and saline solutions through the sense of taste. *J. Psychol.*, 1955, **39**, 157–160.

Eriksen, C. W. Multidimensional stimulus differences and accuracy of discrimination. *USAF., WADC Tech. Rep.*, 1954, No. 54–165.

Eriksen, C. W., & Hake, H. W. Absolute judgments as a function of stimulus range and number of stimulus and response categories. *J. exp. Psychol.*, 1955, **49**, 325–332.

Garner, W. R. An informational analysis of absolute judgments of loudness. *J. exp. Psychol.*, 1953, **46**, 373–380.

Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.

Jones, F. N. The reliability of olfactory thresholds obtained by sniffing. *Amer. J. Psychol.*, 1955, **68**, 289–290.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81–97.

Pollack, I. The information of elementary auditory displays II. *J. acoust. soc. Amer.*, 1953, **25**, 765–769.

(Received July 9, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

THE RELATIONSHIP BETWEEN PROBABILITY DIFFERENCE, $(\pi_1 - \pi_2)$, AND LEARNING RATE IN A CONTINGENT PARTIAL REINFORCEMENT SITUATION

PAUL J. WOODS

Hollins College

In applications of Estes' theory (1950) to two-choice noncontingent reward situations, it has frequently been observed that θ varies directly as a function of π (Restle, 1957). That is, π, the reward probability on the more profitable alternative, an independent variable, is positively related to variation in θ, which is a dependent variable reflecting the rate of learning as well as a theoretical constant representing, in Estes' formulation, the proportion of stimulus elements sampled on any particular trial (Estes & Straughan, 1954). The theory does not predict this variation in θ as a function of π. Yet an important problem is to ascertain just what characteristics of the reinforcement schedules are producing this variation in θ. The present experiment is a step in this direction.

A contingent reward situation was used for several reasons, and here it will be instructive to contrast the contingent situation with the noncontingent. The noncontingent equation for a two-choice situation is as follows (Estes & Straughan, 1954):

$$\bar{p}(n) = \pi - [\pi - \bar{p}(0)](1 - \theta)^n,$$

where $\bar{p}(n)$ represents probability of the more frequently rewarded alternative after n trials. The reward probability is π on one alternative and $(1 - \pi)$ on the other alternative. Hence, referring to π in a two-choice noncontingent situation is the same as referring to the difference between the two probabilities; i.e., these are related in a linear fashion. But as π and the probability difference increase, *S*'s asymptotic performance also increases. In the equation above, θ is a measure of the rate of learning, i.e., the rate at which the fitted exponential curve approaches its asymptote.

In contrast, the contingent equation for a two-choice situation takes the following form (Brand, Sakoda, & Woods, 1957):

$$\bar{p}(n) = \frac{1 - \pi_2}{2 - \pi_1 - \pi_2} - \left[\frac{1 - \pi_2}{2 - \pi_1 - \pi_2} - \bar{p}(0)\right] \times [1 - \theta(2 - \pi_1 - \pi_2)]^n$$

Here π_1 refers to the probability of reward on the more profitable alternative, and π_2 the probability of reward on the less profitable alternative. These probabilities are independent and, hence, their sum may vary from 0 to 2. In contrast to the noncontingent situation, θ is only a component of the term which reflects the rate of learning $(\theta(2 - \pi_1 - \pi_2))$.

Hence, the contingent situation allows greater manipulation of the reinforcement possibilities as well as separation of the stimulus sample proportion from the rate of learning. With increased manipulation over the reinforcement possibilities one is able to hold some things constant while systematically varying others. Specifically, one possibility is to hold

constant the predicted asymptotic probability while systematically varying the probability difference. This was not possible in the noncontingent situation, and we are now able to assess the effects of probability difference on the rate of learning and on θ without the confounding factor of differing asymptotes. The purpose of the present experiment was, therefore, to test for the constancy of θ, and for a direct relationship between probability difference and observed learning rate.

Method

Apparatus.—The two-choice contingent reinforcement apparatus used has been fully described elsewhere (Brand et al., 1957). Briefly, it consists of two levers (microswitches) at table height in front of *S* and a vertical panel containing red and white reinforcing lights and a green signal light. The levers are controlled independently by *E* through a switching device allowing each lever to give either color. By instruction, the red light was selected as a reward.

Procedure.—Five schedules of reinforcement were chosen which yield a common asymptotic prediction of .80 in Estes' prediction equation (1954).

$$p(\infty) = \frac{1 - \pi_2}{2 - \pi_1 - \pi_2}$$

These schedules were 95:80, 90:60, 85:40, 80:20, and 75:0,[1] which have probability differences of .15, .30, .45, .60, and .75, respectively. Ten *Ss* were assigned randomly to each schedule and were run individually for 300 trials.

In an attempt to keep the proportion of responses to the high π alternative fairly low during the first 20 trials, the schedules were manipulated so that the more frequently rewarded alternative was opposite to each *S*'s preferred hand. The *Ss* were then allowed to use only their preferred hands in making responses.

Subjects.—The *Ss* were 50 undergraduate women from courses in psychology at Hollins College who were given a few points extra credit for serving.

Instructions.—When *S* was seated at the apparatus and her preferred hand had been ascertained, the usual instructions were given. Briefly, they consisted of demonstrating the apparatus and telling her that she was to try to outguess *E* and always attempt to get a red light. In addition the following comments were added:

"Now remember, you are going to try to outguess me. The only thing for you to attend to is which switch you pressed and which light flashed. How fast you press, how hard you press, how fast the ready-signal flashes, whether I cough ——. None of these things or anything else but what I've told you enters into the experiment."[2]

Questions were handled by either saying, "I can't tell you that now," or by paraphrasing relevant parts of the instructions.

Results and Discussion

The mean learning curves for each schedule are shown in Fig. 1. Four *Ss* under the 95:80 schedule and one *S* under the 90:60 schedule did not learn to prefer the more rewarded, or high π, alternative in the 300 trials given; mean curves plotted without these *Ss* all appear to approach a common final response level which is somewhat above the theoretical asymptote.

The rate constant can be estimated in various ways (Brand et al., 1957). But the best-fitting curves are obtained by taking empirical starting

[1] The notation used here indicates the probability of the rewarding event, E_1, following each possible response, A_1 and A_2. Thus, for example, 80:40 means that the probability of reward following an A_1 is .80, while the probability of reward following an A_2 is .40. The probabilities were approximated within each block of 20 trials, but for the entire series, the probability schedule was attained.

[2] These statements were used in an attempt to eliminate some of the "hypotheses" *Ss* think up and test when operating in this kind of situation. These statements, as well as the technique of having the more rewarded alternative opposite the *S*'s preferred hand, were directed toward reducing the large amounts of within group variability found in previous research.

points, i.e., the response proportions on the first trial blocks ($\hat{P}(1)$), and asymptotes estimated from the data for each schedule.[3] Values of $\theta(2 - \pi_1 - \pi_2)$ so obtained are plotted in Fig. 2. As can be readily seen, the rate of learning varies directly as a function of the probability difference.

But what about the effect of probability difference on θ itself? If θ remains constant, as the theory predicts, then the functions drawn should be straight lines (illustrated in Fig. 2 for several values of θ). Indeed four of the five estimates of θ shown at the bottom of the graph are almost exactly the same. There is certainly not any consistent variation in θ as a function of probability difference. In particular, there is not the positive relationship that is found in the noncontingent situation.

Here, then, is where the situation stands. Varying the probability difference in the noncontingent situation

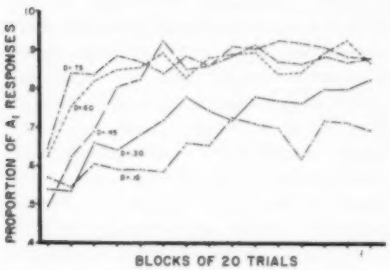

FIG. 1. The mean proportion of responses to the high π alternative in blocks of 20 trials for each schedule.

[3] The final numbers of trial blocks to be averaged in estimating the empirical asymptotes were determined by an inspection of each curve. Following are the number of final blocks averaged and the obtained asymptote for each schedule: 75:0, 7, .881; 80:20, 10, .877; 85:40, 10, .897; 90:60, 3, .808; 95:80, 7, .698.

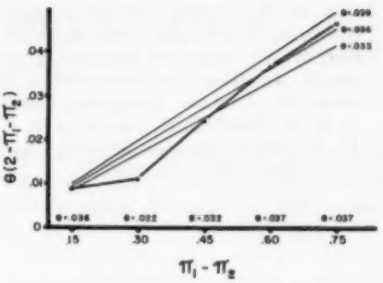

FIG. 2. Empirical values of the rate constant (based on $\hat{P}(1)$ and asymptotes estimated for each schedule) as a function of probability difference.

produces systematic effects on the rate of learning and on the proportion of stimulus elements sampled, both of which are represented by θ. Varying the probability difference in a contingent reinforcement situation while holding asymptotic predictions constant leads to systematic variations in the rate of learning, but θ itself apparently remains constant.

SUMMARY

The purpose of the present experiment was to investigate the effect of probability difference, $(\pi_1 - \pi_2)$, in a contingent reward situation on the rate of learning, represented theoretically by $\theta(2 - \pi_1 - \pi_2)$, and upon the parameter θ alone. Five schedules were used which had probability differences varying from .15 to .75 in increments of .15. According to theory, all schedules yielded asymptotic probabilities of .80 for the more frequently rewarded response. Ten *Ss* were run under each schedule for 300 trials. The results indicate that the rate of learning is directly related to the probability difference while θ apparently remains constant. This result was contrasted with the usual finding in a noncontingent situation where θ varies directly with the probability difference.

REFERENCES

BRAND, H., SAKODA, J. M., & WOODS, P. J. Contingent partial reinforcement and the anticipation of correct alternatives. *J. exp. Psychol.*, 1957, **53**, 417–424.

ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

ESTES, W. K. Individual behavior in uncertain situations: An interpretation in terms of statistical association theory. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes*. New York: Wiley, 1954. Pp. 127–137.

ESTES, W. K., & STRAUGHAN, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

RESTLE, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, **64**, 182–191.

(Received July 9, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

AFTEREFFECT OF A MOVING PATTERN

V. R. CARLSON [1]

National Institute of Mental Health

Visual inspection of a curved line produces an aftereffect of apparent curvature in a subsequently presented straight line. Gibson (1933) originally described this phenomenon in terms of gradual adaptation to the curved line so that it looked more nearly straight, and then gradual recovery from the adaptation during which an objectively straight line would appear curved in the opposite direction. He also observed that the two lines had to be presented approximately in the same place in the visual field, although the effect occurred readily enough without rigidly controlled fixation.

Köhler and Wallach (1944) subsumed these effects of curvature under their more general class of figural aftereffect demonstrations, for which they developed a theoretical explanation in terms of postulated electrophysiological alterations in the cortical tissue corresponding topologically to the locus of the figure. Precisely controlled fixation would not be absolutely necessary, because prolonged general inspection of a figure could build up regions of differentially altered tissue on a statistical basis (Köhler & Fishback, 1950). Osgood and Heyer (1952) offered an alternative explanation based specifically on the occurrence of statistical distributions of excitation corresponding topologically to points and contours in the visual field. Motokawa, Nakagawa, and Kohata (1957) have recently applied the hypothesis of retinal induction to these phenomena, which also depends upon spatially determined relationships though in the retina rather than in the cortical field.

Held (1957) and Köhler (1958) have noted, however, that a pattern which moves across the visual field will also produce a figural aftereffect. To the extent that such a viewing condition produces stimulation distributed more homogeneously over the retina and visual cortex, figural adaptation and aftereffect must necessarily be reduced if they depend fundamentally upon the occurrence of cumulative effects in spatially segregated loci. It is, of course, impossible to produce perfectly homogeneous stimulation and still maintain the perceptual identity of a pattern. If the pattern moves too fast it will begin to blur. The present experiment, therefore, is an attempt, first, to determine whether a clearly established increase in the homogeneity of stimulation brought about by a moving pattern of curved lines, as compared to the more spatially differentiated stimulation of a stationary pattern, will result in a significant decrease in the aftereffect of curvature. Secondly, the motion itself of a moving pattern may be expected to produce aftereffects which must be ruled out in order to determine whether any effect is due specifically to curvature.

METHOD

General procedure.—In the primary experiment a pattern of curved stripes was presented in two main inspection conditions, stationary and moving, with *S* fixating a central stationary point in both cases. On each trial *S*

[1] The author wishes to acknowledge the assistance of Eugene P. Tassone in the collection and analysis of the data.

viewed the inspection pattern for 30 sec. and then immediately adjusted a variable-curvature test line to apparent straightness with stationary fixation at its center. In general, the values for parameters were chosen for experimental or practical reasons other than maximizing aftereffects. Thirty seconds appeared to be about the longest period one could expect most *S*s to maintain steady fixation without exhibiting observably gross eye movements.

Stimulus conditions.—The inspection stimulus-field consisted of a pattern of alternating light and dark stripes, each 10 mm. in width and 145 cm. in radius of curvature. The screen aperture was approximately 57 cm. square, and the pattern completely filled the opening so that the displacement was 28.5 mm. between the midpoint of an arc and the chord joining its ends. This degree of curvature and length of arc were sufficient to make the stripes appear obviously curved, but not great enough to make noticeable the necessarily closer spacing of successive arcs at their ends compared to their centers. The luminance of the light stripes was approximately 30 mL, that of the dark stripes and the surround, 1 mL. The fixation point was a small light spot clearly brighter than the rest of the field.

The test screen was similar to the inspection screen except that the single, dark line was 5 mm. in width and the aperture was 26 cm. square. A radius of curvature of the test line equal to that of the inspection pattern (145 cm.) would therefore produce a displacement of about 6 mm. between the midpoint of the arc and a chord through its ends, and this was approximately the limit of variation of the test line in either direction from straight. For each trial, the test line was preset to a curved position randomly determined in direction and amount. The center of the test screen was located 45° to the right of the center of the inspection screen. Viewing distance for both screens was 105 cm., perpendicular to the center. The test screen was dark during the inspection period and came on as the illumination on the inspection screen went off.

When the pattern was stationary, its precise location with respect to the fixation point was randomly determined for each presentation. This was done specifically in order to achieve a stationary condition as much like the moving condition, on the average, as possible, except in the respect of motion. When the pattern was moving, it progressed steadily across the field at the rate of 24 mm. per sec. This speed was roughly determined to be a satisfactory compromise for very nearly eliminating the brightness afterimage of the pattern, but at the same time not blurring perception of the pattern itself.

The stationary inspection pattern was varied in horizontal-vertical orientation and in direction of curvature, comprising four separate conditions: stripes horizontal, curved convex down or convex up; stripes vertical, convex to the right or convex to the left. The moving pattern was varied in orientation, direction of curvature, and direction of motion, making eight separate conditions—for example, stripes horizontal, curved convex up, moving down. The test line was always oriented horizontally if the inspection pattern was horizontal, vertically if the inspection pattern was vertical. These variations were used (*a*) to eliminate general cumulative or habituative effects which might occur if the pattern were always presented in the same way, and (*b*) to counterbalance effects of specific combinations of orientation, direction of curvature, and direction of motion.

In addition, a control condition was included in which *S* looked at the fixation point without an inspection pattern present, and another control condition in which the inspection pattern was a pattern of straight stripes, of the same dimensions as the curved ones, oriented horizontally or vertically and moving up, down, right, or left.

These various inspection conditions were arranged in blocks of trials such that horizontal conditions alternated with vertical conditions, each specific condition occurred once in each block, and the same condition never followed the same preceding condition in different blocks. The *S* completed 2 blocks of trials on the first testing day, 3 on the second day, and 3 on the third day. At least several days, usually one week, elapsed between successive testing days. Two minutes elapsed between successive trials.

Apparatus.—The inspection pattern was projected onto a translucent plastic rear-projection screen by means of a modified 35-mm. filmstrip projector, equipped with a dove prism so that the projected image could be rotated to any orientation on the screen. For the moving pattern the film was pulled through the film gate by a motor drive with the shutter continuously open. The test line consisted of a curved rod mounted immediately behind a similar plastic screen, illuminated so that *S* saw the shadow cast by the rod. As he turned the control knob the rod rotated around a longitudinal axis such that the ends of the (shadow) test line curved toward or away from straightness while the center remained stationary. Settings were read from

a scale indicating rotational position of the rod in degrees, which were converted to millimeter displacement between the center of the test line and a chord through its ends.

Instructions.—The *S* was instructed to concentrate carefully on the fixation point, not to look away from it, and not to let his eyes wander away from it. A ready signal occurred a few seconds before the end of the inspection period, but *S* was cautioned not to look away until the light actually went off. He was told to shift quickly from the inspection screen to the test screen and to adjust the test line quickly so that it appeared about straight. Four seconds were allowed, but with practice *S* could easily make the adjustment within 2 sec. One other point was stressed, namely, that *S* take care not to tilt his head, especially in shifting from the inspection screen to the test screen.

Observations.—The *E* observed *S*'s eyes during each trial, but the primary criterion for adequate fixation was *S*'s report of afterimages. If *S* maintained stationary fixation, he should report an afterimage of brightness contrast when the inspection pattern was stationary but not when it was moving. The *S*'s report was elicited as informally as possible in terms of periodic general discussions as to whether the 2-min. interval between trials was sufficient to allow for dissipation of any afterimages which might occur. It was never suggested to *S* that he should or should not see an afterimage following one condition rather than another.

Subjects.—The *S*s were 10 paid college student volunteers, 6 men and 4 women. Three were psychology students but were not acquainted with the subject matter of this investigation. Ten additional *S*s from the same population were used in a further control experiment. All had normal or corrected vision.

Secondary control experiments.—Five of the *S*s who had served in the main experiment—those still available—were subsequently tested in a repetition of the experiment using an altered aperture on the inspection screen. The original aperture was made relatively large to minimize possible effects of the aperture itself. But with any aperture not curved at the edges to correspond to the curvature of the pattern, a moving pattern of curved lines produces an impression of motion as the lines intersect the edge of the aperture. This apparent motion occurs at right angles to the direction of travel of the pattern, toward or away from the center of the edge depending on the direction of curvature and the direction of travel. To check whether this effect might influence the aftereffects obtained, an aperture was used in which the two relevant edges were curved to the same degree as the pattern.

A comparison was also made between the stationary curved pattern and a stationary single dark line of the same curvature, both used as inspection conditions in the manner already described. This single line was of the same width as the test line (5 mm.), and fixation was at its center. Ten additional *S*s and the above-mentioned five were used in this comparison.

Body sway suggestibility scores were available on the first 10 *S*s. These were obtained primarily for another purpose, but a variable such as suggestibility seems to the writer particularly relevant to the aftereffect kind of experiment. Although *S* may not surmise the specific hypothesis involved, the situation obviously has to do with the influence of a preceding stimulus on subsequent perception of the test stimulus. The results may very well be biased by how *S* may think the test stimulus "should" be affected. For the body sway test, a device tracing a continuous curve of forward and backward movement was connected to a belt around *S*'s waist. The *S* stood with eyes closed and feet together, while a tape-recorded voice suggested that *S* was falling forward. The test period lasted 2 min. and was preceded and followed by a ¾-min. control period with no suggestion. The measure was the mean height of the curve above the baseline value.

Results

Preliminary considerations.—All of the scores have been computed so that a positive value reflects an aftereffect in the predicted direction irrespective of the orientation of the inspection pattern. This direction has been well-established by previous research: After viewing a curved inspection line, a straight test line will appear curved in the opposite direction; *S* must therefore adjust the test line in the same direction of curvature as the inspection curve in order for the test line to appear straight. In the present data a score of zero means no aftereffect, and a negative value indicates an adjustment opposite in direction to the predicted direction.

In general each separate condition produced a statistically significant effect in the positive direction. For this test the stationary pattern and the moving straight lines were compared with the control condition of no inspection pattern, and the moving curved pattern was compared with the moving straight lines as a control condition. It was possible to combine the separate horizontally-oriented conditions together and the vertically-oriented conditions together. The vertical conditions tended, however, to show a greater effect than the horizontal conditions. This difference was not of clear statistical significance, but it occurs regularly throughout the data. Secondly, the within-*Ss* variances for the horizontal conditions, including the control settings, were generally smaller than those for the vertical conditions. Horizontal-vertical differences in figural aftereffects have sometimes been reported in the literature. But in the present situation it is just as possible that *S* could adjust a horizontal line to objective straightness more accurately than he could a vertical line, allowing less opportunity for the aftereffect to manifest itself. There was also a difference in relationship to the body sway test. The correlations of the suggestibility scores with the aftereffect of the stationary pattern and the moving horizontal pattern were positive but low (average $r = .32, P > .10$). The only significant correlation was with the aftereffect of the moving vertical pattern ($r = .68, P < .05$). Why this particular condition should exhibit a relationship more so than the others is not clear, and the finding may be a chance one. It is possible, however, that the aftereffect for the moving vertical pattern may have been artificially enhanced by some attitude *S* tended to take toward it. In any event, for these several reasons, the results are presented separately for the horizontal and vertical conditions.

There were no significant differences due to testing days, nor any demonstrable trend over the three testing days. The reliability coefficients for individual *Ss*' scores were of borderline significance. Probably the reason is that an individual's score for a single day is based on too few determinations. Validity of the measurement of aftereffect here depends on eliminating certain error factors, such as initial position of the test line, by randomizing the direction of their effects, and they very well might not be canceled out in a limited number of trials.

The results using an aperture with two curved edges, eliminating apparent transverse motion, suggested no essential difference due to this factor. This aperture by itself, without an inspection pattern present, produced no aftereffect, and the aftereffects of the inspection patterns were approximately the same as those obtained using the aperture with straight edges. With a smaller aperture, apparent motion effects at the edges might, of course, have some influence.

Eye movements.—As far as could be observed by *E*, the *Ss* maintained very well-controlled stationary fixation. Occasional flicks of the eyes away from the fixation point and back to it again occurred, but it was not possible to differentiate between the stationary and moving stimulus conditions by watching *S*'s eyes. In particular, no following movements were noted, such as would occur if *S* fixated a moving point periodically. Somewhat more frequent eye movement occurred during adjustment of the test line. This would seem reasonable, for here *S* was actively con-

cerned with the curvature of the line whereas in the inspection condition the lines were merely present. Some *Ss*, in fact, said they more or less ignored the pattern and could not, with any assurance, report what particular inspection condition had been presented after a trial had been completed. One *S* maintained good fixation but was unable to refrain from tilting his head recurrently. This particular individual showed some average aftereffect for the stationary pattern and none for the moving pattern.

The *Ss* were unanimous in indicating that they experienced a clear and persisting brightness-contrast afterimage of the pattern when it was stationary, but only a vague and short-lived afterimage, or none at all, when the pattern was moving. An afterimage of the stationary single line was detectable, but it was not nearly as obvious or persistent as that of the stationary pattern. These afterimage effects were reported by the *Ss* in terms of looking at a darkened, homogeneous surface between trials. They indicated that even the afterimage of the stationary pattern was not particularly noticeable against the illuminated test screen during the test period. Occasional comments apparently referred to a motion aftereffect (waterfall illusion), but these were rather indefinite.

It cannot be claimed that fixation was precisely the same for the stationary and the moving conditions, nor that the small, rapid eye movements which are always present were necessarily the same for both conditions. It seems altogether certain, however, that stimulation by the moving pattern was spatially distributed over the retinal and cortical fields to a greater degree than was the case for the stationary pattern.

Effect of motion.—The moving pattern of *straight* lines produced an aftereffect of apparent curvature on the test line. Except for smaller magnitude, this aftereffect of motion in a given direction was similar in result to the aftereffect produced by a stationary curve that is curved concavely in the same direction. For example, in adjusting the test line to apparent straightness, *S* tended to produce a curve concave on the right after viewing the straight line pattern moving to the right, and also after viewing the stationary pattern of curves, curved concave on the right. An aftereffect of the moving curved lines may therefore be looked upon as compounded of an aftereffect due to curvature and an aftereffect due to motion. If these two sources of effect are separable, then the aftereffect of the stationary curved lines plus the aftereffect of the moving straight lines should add up to an estimate of the aftereffect obtained from the moving curved pattern when it is concave in the direction of its motion. And the aftereffect of the stationary curves minus the aftereffect of the moving straight lines should approximate the aftereffect produced by the moving curved pattern when it is convex in the direction of its motion. The averages do agree reasonably well with just such a simple additive hypothesis (Table 1). The estimates based on the combination of stationary curvature and moving straight lines do not differ significantly from the values obtained with the moving curves.

When the two moving curved conditions concave in the direction of motion and convex in the direction of motion, are averaged together, then the effect of motion should be balanced out leaving a measure reflecting the aftereffect due to curvature, only, of the

TABLE 1

EFFECT OF MOVING CURVES COMPARED WITH THE SUM AND DIFFERENCE OF THE EFFECTS OF STATIONARY CURVES AND MOVING STRAIGHT LINES[a]

Orientation	(A) Curved Lines Stationary	(B) Straight Lines Moving	(A) + (B)	Curved Lines Moving Concave in Direction of Motion	(A) − (B)	Curved Lines Moving Convex in Direction of Motion
Horizontal						
Mean	.58	.22	.80	.63	.36	.25
SD	.34	.14	.31	.40	.41	.20
P[b]	<.001	<.001	<.001	<.01	<.05	<.01
Vertical						
Mean	.70	.38	1.08	.93	.33	.38
SD	.35	.41	.59	.48	.48	.32
P[b]	<.001	<.05	<.001	<.001	<.10	<.01

[a] Opposite directions of motion averaged together to obtain displacement from straight (mm.).
[b] Probability associated with the *t* value for the difference of the mean from zero, $df = 9$.

moving pattern. This expectation is supported by the comparisons in Table 2, which exhibit no significant effect of direction of motion in the final averaged values for the moving curved pattern. The scores for some individual *S*s agree well with the average results of this analysis, but there are sizeable discrepancies for others.

Stationary vs. moving conditions.—under the presumption, then, that the final averages for the moving curved pattern have been freed from the influence of motion itself, these averages may be compared with those for the stationary curved pattern (Table 3). The Stationary means are based on 16 determinations for each *S*, and the Moving means are based on 32 determinations. The moving pattern did not produce a clear decrement in aftereffect. As mentioned previously, the Vertical Moving value might have been raised due to some possible suggestive influence in the situation, leading to an artificially small difference here. But neither is the Horizontal Moving value significantly less than its corresponding Stationary value ($P > .10$).

Magnitude of effect.—The over-all level of effect here amounts to 2′ of visual angle. This order of magnitude falls at the lower end of the range of values reported in a recent summary of figural aftereffect experiments by various investigators (Sagara & Oyama, Table 1). The parameters of the present experimental conditions were not chosen primarily to produce maximal effects. The one most probably not optimal was sharpness of the contours, which was necessarily reduced somewhat by the use of a rear-projection screen. On the other hand, the use of a pattern of curves might itself have reduced

TABLE 2

COMPARISON OF OPPOSITE DIRECTIONS OF MOTION FOR MOVING CURVES[a]

Meas.	Horizontal		Vertical	
	Moving Up	Moving Down	Moving Right	Moving Left
Mean	.42	.46	.60	.70
SD	.30	.30	.24	.31
p[b]	<.01	<.01	<.001	<.001

[a] Opposite directions of curvature averaged together to obtain displacement from straight (mm.).
[b] Probability associated with the *t* value for the difference of the mean from zero, $df = 9$.

the aftereffect, so that it was pertinent to determine the effect of a single, isolated line, presented the same way under the same circumstances. A preliminary check using 5 of the *Ss* in the above experiment suggested the possibility of such a difference. The average effect was .88 mm. for the single curved line.

In order to obtain a sufficiently large *N* for a statistical test the sample was augmented by recruiting 10 additional *Ss*. The results for all 15 *Ss* are summarized in Table 4. The pattern may not produce quite as great an aftereffect as a single curve, but the general order of magnitude appears to be the same. Gibson (1933) came to the same conclusion in comparing a single curved line with a pattern of curves. The greater difference obtained with the first 5 *Ss* was most likely a matter of sampling fluctuation.

Discussion

The aftereffect of motion itself in producing apparent curvature is probably a direct consequence of the fact that effects such as the waterfall illusion are greater toward the periphery of the visual field than in the center (Wohlgemuth, 1911, pp. 109–110). Smith (1952) has criticized theories of figural aftereffects for not being able to account for the aftereffects of perceived motion, and George (1953) and Deutsch (1956) have made certain suggestions which might lead to theoretical conceptualizations encompassing both kinds of effect. At the present time the only reason for believing that similar causes are involved is that the end effects are at least superficially similar. The results obtained here indicate that these effects are separable, and the following discussion presumes that the influence of motion, as such, has been eliminated. This experiment has no crucial bearing on whether figural aftereffects and the aftereffects of seen movement are due to the same or different fundamental processes.

The fact that the curvature of the inspection pattern produced an aftereffect, whether it was stationary or moving, would appear to suggest that topological representation of the contours in the visual system is not critically important for the aftereffect to occur. There is the possibility that a more general kind of perceptual normalization takes place which may not be altogether localized in the retina or the visual cortex. Gibson (1933, 1937) partly placed this kind of interpretation upon the aftereffects of curvature and tilted lines. Köhler and Wallach (1944, pp. 311–312) agreed that such normalization may occur but argued that it represents a class of phenomena different from figural aftereffects and specifically does not explain the aftereffects of curvature and tilted lines. Using rectangles as stimuli, Fox (1951) also obtained experimental

TABLE 3

Moving Curves Compared with Stationary Curves[a]

Meas.	Horizontal		Vertical	
	Stationary	Moving	Stationary	Moving
Mean	.58	.44	.70	.65
SD	.34	.23	.35	.26
p^b	<.001	<.001	<.001	<.001

[a] Opposite directions of curvature and motion averaged together to obtain displacement from straight (mm.).
[b] Probability associated with the *t* value for the difference of the mean from zero, $df = 9$.

TABLE 4

Stationary Curved Pattern Compared with Stationary Single Curved Line[a]

Meas.	Horizontal		Vertical	
	Single Line	Pattern	Single Line	Pattern
Mean	.73	.67	.86	.77
SD	.18	.33	.49	.35
p^b	<.001	<.001	<.001	<.001

[a] Opposite directions of curvature averaged together to obtain displacement from straight (mm.).
[b] Probability associated with the *t* value for the difference of the mean from zero, $df = 14$.

results consistent with the view that two different kinds of adaptation may be involved. Nozawa (1953) found that the aftereffect of curvature of a preceding inspection line was always in the direction of changing a curved test line so that it appeared more nearly straight, even under circumstances where analogy to figural aftereffects would predict an increase in curvature. Related to these studies in the matter of spatial localization is the uncertainty over whether apparent perceptual size or retinal image size, or both, determines the aftereffect (Sutherland, 1954). In these various experiments, however, it is clear that whatever the specific consequences of the particular spatial arrangement of the stimuli, the adaptation and aftereffect are at least approximately localized to that part of the visual field where the inspection stimulus is presented.

In accounting for the aftereffect of moving curves it is possible to retain the requirement that the inspection stimulus be localized in the visual field if we abandon the necessity of a cumulative build-up in adaptation over time. When the stimulus pattern is terminated it does leave a localized pattern of stimulation wherever it happened to be at that particular moment. This is not an *ad hoc* hypothesis by any means. Köhler and Wallach (1944, pp. 354–355) observed that in some instances evidence of figural aftereffects could be obtained with inspection times of less than a second. Oyama (1953) and Ikeda and Obonai (1953) found that lengthening the inspection period produced no clearly greater aftereffect but decreased the rate of its decay. They point out that if the aftereffect is not measured immediately after the end of the inspection period, the magnitude of the aftereffect will appear to increase as a function of lengthening the inspection time. In reality its persistence, not magnitude, may be a function of inspection time. Further, perceptual alternation in ambiguous figures has been related to figural aftereffects (Carlson, 1953; Hochberg, 1950). If the same process underlies such alternation, the adaptation must develop fairly quickly in order to produce the rate of alternation normally encountered. The generality of this formulation has yet to be established, and in particular its applicability to the Köhler-Wallach figural aftereffects may be questionable.

At this point it does not seem very satisfactory to postulate a spatial projection of time to the brain as Köhler (1958) suggests. Such an hypothesis appears not to be necessary to account for the aftereffect of moving curves, nor, possibly, the aftereffects of scanning a pattern (Held, 1957). On the other hand, the hypothesis does not make it any easier to unify many of the diverse effects which are similar yet not clearly the same as figural aftereffects (Sagara & Oyama, 1957; Spitz, 1958). More serious consideration needs, rather, to be given to the possibility that central processes other than those localized in the primary projection areas are also intimately involved in visual adaptation and aftereffect.

Summary

The aftereffect of a pattern of curved lines was determined when the pattern was stationary and when it was moving. Fixation was stationary in both cases, and the test for aftereffect was performed immediately after viewing the pattern. It was possible to balance out the aftereffects of motion and obtain an estimate of effect due to curvature only. The aftereffect of curvature was not decreased to a significant degree by moving the pattern across the visual field. This result may be accounted for in terms of the topological representation of contours in the visual cortex, if it is assumed that momentary adaptation produces as great an aftereffect as longer periods of adaptation when the aftereffect is measured immediately.

References

Carlson, V. R. Satiation in a reversible perspective figure. *J. exp. Psychol.*, 1953, **45**, 442–448.

Deutsch, J. A. The statistical theory of figural after-effects and acuity. *Brit. J. Psychol.*, 1956, **47**, 208–215.

FOX, B. H. Figural after-effects: "Satiation" and adaptation. *J. exp. Psychol.*, 1951, **42**, 317–326.

GEORGE, F. H. On the theory of the figural after-effect. *Canad. J. Psychol.*, 1953, **7**, 167–171.

GIBSON, J. J. Adaptation, after-effect and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, **16**, 1–31.

GIBSON, J. J. Adaptation with negative after-effect. *Psychol. Rev.*, 1937, **44**, 222–243.

HELD, R. Patterning versus segregation in the production of visual after-effects. Paper read at East. Psychol. Ass., New York, April, 1957.

HOCHBERG, J. E. Figure-ground reversal as a function of visual satiation. *J. exp. Psychol.*, 1950, **40**, 682–686.

IKEDA, H., & OBONAI, T. The quantitative analysis of figural after-effects: I. The process of growth and decay of figural after-effects. *Jap. J. Psychol.*, 1953, **23**, 246–260; **24**, 59–66, 84–85.

KÖHLER, W. The present situation in brain physiology. *Amer. Psychologist*, 1958, **13**, 150–154.

KÖHLER, W., & FISHBACK, JULIA. The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories. *J. exp. Psychol.*, 1950, **40**, 267–281.

KÖHLER, W., & WALLACH, H. Figural after-effects; An investigation of visual processes. *Proc. Amer. phil. Soc.*, 1944, **88**, 269–357.

MOTOKAWA, K., NAKAGAWA, D., & KOHATA, T. Figural after-effects and retinal induction. *J. gen. Psychol.*, 1957, **57**, 121–135.

NOZAWA, S. Prolonged inspection of a figure and the after-effect thereof. *Jap. J. Psychol.*, 1953, **23**, 217–234; **24**, 85–87.

OSGOOD, C. E., & HEYER, A. W., JR. A new interpretation of figural after-effects. *Psychol. Rev.*, 1952, **59**, 98–118.

OYAMA, T. Experimental studies of figural after-effects: I. Temporal factors. *Jap. J. Psychol.*, 1953, **23**, 239–245.

SAGARA, M., & OYAMA, T. Experimental studies on figural after-effects in Japan. *Psychol. Bull.*, 1957, **54**, 327–338.

SMITH, K. The statistical theory of the figural after-effect. *Psychol. Rev.*, 1952, **59**, 401–402.

SPITZ, H. H. The present status of the Köhler-Wallach theory of satiation. *Psychol. Bull.*, 1958, **55**, 1–28.

SUTHERLAND, N. S. Figural after-effects, retinal size and apparent size. *Quart. J. exp. Psychol.*, 1954, **6**, 35–44.

WOHLGEMUTH, A. On the after-effect of seen movement. *Brit. J. Phychol.* Monogr. Suppl., 1911, No. 1, 1–117.

(Received July 24, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

MOTION PARALLAX AS A DETERMINANT OF PERCEIVED DEPTH[1]

ELEANOR J. GIBSON, JAMES J. GIBSON, OLIN W. SMITH, AND HOWARD FLOCK

Cornell University

Motion parallax is the optical change of the visual field of an observer which results from a change of his viewing position. It is often defined as the set of "apparent motions" of stationary objects which arise during locomotion. Psychologists assert that it is a "cue" for perceiving the depth of the objects, but the optical fact of motion parallax must be distinguished from its capacity to induce perceptions. It has not been experimentally demonstrated that motions in the field of view will actually yield corresponding judgments of depth. This is a purely psychological problem. The optics of motion parallax, on the other hand, is a problem for geometry and ecology.

Recently, the suggestion has been made that a continuous gradient of motions in the field of view will induce the perception of slant-depth (J. J. Gibson, Olum, & Rosenblatt, 1955) inasmuch as the perception of depth is intimately connected with the perception of surfaces (J. J. Gibson, 1950). This statement also needs experimental test. The purpose of the present study is to investigate what kinds of motion in the light entering an eye do in fact consistently arouse certain judgments of depth, and what do not.

The experiments must be carried out with artificial motions in a field of view rather than those obtained in a natural environment if we wish to study the effect of motion parallax in isolation from other cues or stimuli for depth. The variables of size, density, linear perspective, differential blur, and binocular parallax should be eliminated or so reduced as to be ineffective in the array of light entering *O*'s eye. A method of achieving this result has been devised, and a suitable control employed.

The experimental method should also preclude actual movement or locomotion of *O*. If the cue of motion parallax is so defined as to require *active* head movement or locomotion, proprioceptive and vestibular stimulation is also present. This definition is unjustified, since passive locomotion in trains and airplanes should be admitted as circumstances when motion parallax occurs. Certain patterns of motion in the field of view of *O* do induce impressions of being moved through space if we accept as evidence the illusions of locomotion obtained in viewing a panoramic motion picture, or in a training device for simulating aerial flight.

Perception of absolute distance and of relative depth.—The apparent displacements of the sensations of objects are said to be cues for perceptions of their depth. What kind of depth? The question arises whether their distances from the perceiver can be judged or whether they will only appear to be separated in the third dimension. Helmholtz, in his description of motion parallax, asserted both hypotheses. On one page he described the appearance of objects "gliding past us" as we walk through

[1] This work was supported by the Office of Naval Research under Contract NONR 401 (14) with Cornell University. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

the countryside, and asserted that "evidently under these circumstances the apparent angular velocities of objects in the field of view will be inversely proportional to their real distances away; and consequently safe conclusions can be drawn as to the real distance of the body from its apparent angular velocity" (Helmholtz, 1925, p. 295). On the next page he described the appearance of an indistinguishable tangle of foliage and branches in a thick woods as a man stands motionless, but noted that "the moment he begins to move forward, everything disentangles itself and immediately he gets an apperception of the material contents of the woods and their relations to each other in space, just as if he were looking at a good stereoscopic view of it" (Helmholtz, 1925 p. 296).

In the first quotation Helmholtz says that angular velocity is a cue for the perception of absolute distance. In the second, he suggests that a difference in angular velocity is a cue for the perception of separation in depth, or of relative distance only. These two hypotheses are by no means the same, and they should be considered separately and tested separately. We will be concerned here primarily with the second.

Two-velocity motion parallax and flow-velocity motion parallax.—Although motion parallax has been said to apply to the whole array of objects in an environment and a large array of apparent motions in the field, the experiments performed have in the past been confined to two objects and two velocities in a restricted field of view.

Bourdon (1902) reported experiments in which *O* looked with one eye at a pair of luminous spots in a dark corridor. The sources were at different distances but the spots were of the same angular size. When the head was fixed with a biting-board, *O* "could not judge at all accurately" which light was the nearer, but with the slightest movement of the head from side to side "it was easy to judge" the relative depth of the two. But the absolute distance of neither light was detectable.

Tschermak-Seysenegg (1939) improved on this arrangement with what he called a "parallactoscope," by analogy with the stereoscope. He defined motion parallax as arising from movement either of a group of visible objects on the one hand, or the position of *O*'s eye on the other, emphasizing the relativity of the situation. But, he studied only the detection of depth of two objects with voluntary head movement. His apparatus was a modification of the familiar two-pins setup used to obtain the threshold for binocular depth perception. It permitted *O* to move one eye from side to side, with a sliding headrest, so as to obtain *successive* impressions equivalent to the *simultaneous* impressions obtained with both eyes open. The average error of equating the distance of the vertical wires was small under these conditions, although not as small as the error with both eyes open. When only one eye with a fixed head was used, the error was very large.

Graham, Baker, Hecht, and Lloyd (1948; see also Graham, 1951) obtained the threshold for separation in depth of two needles pointing toward one another, as seen on a uniform field through a window. The needles moved from side to side on a common carriage. They appeared to be aligned at the center of their motion cycle and offset at the extremes of the cycle unless the adjustable needle had been set into the same frontal plane as the fixed needle. Graham thus eliminated for the first time in this type of experiment the additional sensory information produced by voluntary head movement.

More exactly, what Graham obtained was the just noticeable difference between two angular velocities in a field of view, under probably optimal conditions. The threshold was extremely low—about 30 sec. of arc per second of time. It is notable that the reports of what *O*s perceived, however, were not unanimous. Some saw the separation in depth as such; others perceived either the difference in velocity of the two needles or noticed the change of alignment or offset of the needles. Although

the latter impressions may be cues for the former, the experiment was not concerned with the effectiveness of such cues for producing depth impressions.

Somewhat later, a case of motion parallax different from the two-velocity case was defined mathematically by Gibson, Olum, and Rosenblatt (1955). This is the array of angular velocities of the optical elements projected from a surface to a moving station point. There is a flow of velocities rather than a pair of velocities in such an array, and the phenomenon was named motion perspective to distinguish it from motion parallax as it had been studied up to that time.

For the study of the perceptions induced by flow-velocity in a field of view, including gradients of velocity, skew motions, and transformations, a different sort of apparatus is required from that previously employed. The two-velocity experimenters used a pair of real objects at real distances to produce the optical motions. More freedom is achieved by using a projection screen or some other optical means to produce them. The experiments to be described used shadows on a translucent screen. Accommodation is thereby controlled.

Several exploratory experiments have been published on flowing motion. They are of various types, and they have been produced in various ways. J. J. Gibson and Carel (1952) attempted to induce the perception of a receding surface in a darkroom with a bank of luminous points which carried a gradient of velocities. This stimulus failed to arouse the perception of a surface, however, and the depth judgments were ambiguous. O. W. Smith and P. C. Smith (1957) investigated the perception of convexity or curvature of a textured surface with various combinations of depth cues, including the flow-velocity type of motion parallax. Although motion in the field contributed to the judgment of convexity, in no case did motion cause a surface otherwise judged as flat to be judged as curved. Hochberg and O. W. Smith (1955) studied the perception of depth induced by the centrifugal flow of luminous pattern elements in the dark, the expansion phenomenon. J. J. Gibson and E. J. Gibson (1957) investigated the perception of the rigid rotation of an apparent surface elicited by the continuous perspective transformation of regular and irregular patterns or forms.

These experiments differed in the structure of the optic array used to "carry" the motion in question, and they also differed in the degree to which perceptions of space were aroused. They led to the choice of the kind of random texture employed in the present experiments, which is intended to yield the experience of a plane surface.

What is now needed is an experimental comparison of the judgments obtained with *two* velocities in a field of view and those obtained with *many* velocities in a field of view. Although no clear line can be drawn between them, the two-velocity type of motion parallax applies to the problem of perceiving a group of objects in otherwise empty space, while the flow-velocity type of motion parallax applies to the perceiving of a background surface such as a wall (or substratum). These are not the same problem for perception even though it may be difficult to distinguish sharply between their respective kinds of stimulation.

Optical geometry of motion parallax.— The environmental situation which leads to an array of different motions in a visual field should be defined more carefully. This is the optical geometry of motion parallax, as distinguished from the visual appearance of motion parallax. Graham (1951, pp. 878 ff.) has given the geometry of certain special cases of this situation. J. J. Gibson et al. (1955) have analysed the case of an extended surface such as the ground. What will be discussed here is the case of an environment of discrete objects.

When light rays from permanent objects of an environment converge to a point, they constitute what may be called an optic array, and the elements of this array constitute a pattern. An eye or a camera at the station point can register this pattern of luminous elements. If the point moves, the

pattern is altered in a way which depends on both the displacement of the point and the layout of the objects. How the eye responds to this alteration of pattern is our problem.

The first question is how to specify mathematically the change of pattern in a way that is relevant for vision. By choosing a coordinate system for the array, one can specify the absolute position of each element and the displacement of each element per unit of time, that is, its absolute angular velocity. It would then be true in a certain sense, as Helmholtz (1925, p. 295) said, that "the apparent angular velocities of objects in the field of view will be inversely proportional to their real distances away." But more exactly, it would be true only if the linear velocity of the station point were constant (J. J. Gibson et al., 1955). A given angular velocity is a cue for distance, or permits a "safe conclusion" about distance, only if the speed and direction of one's locomotion is known.

The trouble with positions and angular velocities of elements in a field is the difficulty of understanding how an eye can register them. As the Gestalt theorists have emphasized, what the eye seems to pick up is the mutual separation of elements, their pattern, rather than their position or directions. And, accordingly, it is easier to suppose that the eye responds to changes of separation, or change of pattern, rather than to absolute displacements or velocities. Helmholtz might better have asserted that a difference between the angular velocities of two elements in the field will be directly related to the difference in distance between the corresponding objects in space. Such relative velocities involve a transformation of pattern, and this may be what the eye is primarily sensitive to. It is not immediately evident what the best method is for specifying the information about objects in an array of light projected to an eye.

But one fact should be clear. Only if there is an eye at the point of projection and only if it is sensitive to the motions in the optic array, relative or absolute, does a psychological question arise. Will the possessor of the eye see merely the change of pattern of the array? Or will he see moving objects in the field of view? Or will he see stationary objects at different distances? In order to show that motion parallax is effective for the perception of depth it must be demonstrated experimentally that differential motions in an array of light to an eye will yield differential judgments of depth. And the array should be such that when the motion is eliminated the judgments of depth will cease, for only then will motion parallax have been isolated from other cues for depth.

Experiment I: Motion Parallax With Two Velocities

Problem and Method

The two-velocity experiments were repeated with (*a*) two spots in a field to carry the motions, and (*b*) two superimposed textures filling the field to carry them. In both cases the velocity *difference* was taken to be the essential cue for possible judgments of depth, not the absolute velocities. In this experiment, reports were obtained for a large velocity difference, a small velocity difference, and no velocity difference, that is, a motionless field. The last was a control.

Apparatus and stimuli.—The light entering *O*'s eye came from the translucent screen of a point source shadow-projector (J. J. Gibson, 1957; J. J. Gibson & E. J. Gibson, 1957). He saw only a luminous rectangular field in which dark circles or textures could be made to appear and to move. These were actually the shadows of opaque substances attached to a transparent mount behind the screen. This was a large sheet of glass or plastic whose edges were never visible. Differential translatory velocity of the shadows was produced with two mounts, one behind the other, which could be made to move parallel to the screen on a common carriage. The array of light to the eye was simply the reverse of the array projected to the screen, since the eye and the point source were symmetrically located equidistant from the screen (Fig. 1). The window was 32.2 × 36 cm. at a distance of 126 cm. from the eye, subtending an array 14.5° high and 16.2° wide. The window was viewed through an aperture by a seated *O*.

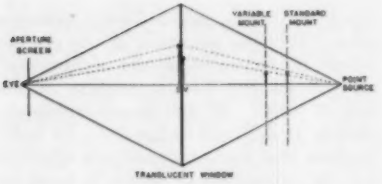

FIG. 1. The shadow projector viewed from above. In a unit of time, the shadow of a spot at the center of the standard mount sweeps through a certain angle and that of a corresponding spot on the variable mount sweeps through a lesser angle, as shown. The two mounts roll on the same carriage. If they are close together, there is no difference in angular velocity, but as the variable mount is positioned farther from the point source and closer to the screen, the angular velocity of its shadow decreases. With this apparatus, it can decrease to about one half of the angular velocity of the standard. By trigonometry, the ratio of the lesser (V) to the greater (S) angular velocity is equal to the inverse of the ratio of the distances of their respective mounts from the point source. In the diagram above, it is about 0.7.

The carriage which bore the two mounts rolled silently on tracks and could be pulled from side to side through an excursion of 45 cm. It was operated by hand to produce a motion cycle in about 8 sec. A small shutter close to the point source enabled *E* to eliminate the shadow between trials, leaving the screen illuminated by diffuse light.

The two adjacent spots in the field were produced by attaching small paper circles to each mount, at different elevations so that their shadows did not pass through one another as they moved across the field. The faster spot was above the slower spot. The diameter of both was 5.2°, one paper circle being compensated in size to match the shadow of the other.

The superimposed random textures were produced by a technique of sprinkling talcum powder over the surfaces of the two transparent mounts. This yields an optical texture with indefinite contours and indefinite elements. When the two were superimposed but motionless, they constituted a single texture with no cue for superposition, and gave the appearance of a single surface, something like that of a cloud. This apparent surface filled the whole window and appeared at an indefinite distance from *O*.

As noted, the two angular velocities as such were not uniform, decreasing to zero at either end of a motion cycle, and changing direction alternately. Minor variations in velocity also occurred as a consequence of moving the carriage by hand. The independent variable of this experiment was the difference in velocity between the two shadows. It was expressed as the ratio of the slower (the variable) velocity to that of the faster (the standard) velocity, or V/S.

Procedure.—Each *O* was seated at the apparatus, asked to apply his preferred eye to the aperture, and instructed simply to "describe what he saw in the window." He was first presented with a motionless field for as long as he needed to make a report, which was recorded. He was then presented with continuous cycles of motion at the maximum velocity difference (V/S = .51) until his report was completed. Finally he was given the minimum velocity-difference (V/S = .97). The *E* made no comment at any time, since wholly spontaneous reports were desired. The order of presentation was intended to minimize the effect of suggestion on the perceiving of depth.

A group of 26 *O*s went through this procedure with the spot field and another group of 46 with the textured field. Formal judgments and answers to questions were obtained afterwards from some *O*s, which will be described when relevant. They were requested in the terms used spontaneously by the *O*.

Results

The words used by the *O*s to describe what they saw varied widely, and the effect to identify things was reminiscent of descriptions of cognitive inference (Vernon, 1957). But the reports could later be classified easily with respect to depth or distance. The motionless textured field was unanimously reported to be a single surface without any difference in depth. The motionless spots, however, were reported at different distances by 4 of the 26 *O*s. The spots, therefore, did not wholly satisfy the requirement that impressions of depth be absent in the absence of motion, although the combined textures did.

A large velocity difference (.51) for the textured field always gave a

perception of two surfaces separated in depth, as evidenced by the reports of all 46 *O*s. For the spot field, the reports were not unanimous, but 22 out of 26 *O*s did describe a difference in depth of the two objects.

The small velocity difference (.97) was evidently close to the threshold. None of the *O*s reported two separated surfaces for the textured field, and only 7 out of 26 reported different distances for the spots.

The direction of the difference in depth reported was not unanimous for either the spots or the textures. Insofar as two-velocity parallax is a reliable and effective indicator of relative depth, the faster velocity should correspond to the nearer object or surface. But 7 out of 26 *O*s saw the slower spot as the nearer object instead of the reverse, and 10 out of 46 *O*s saw the slower texture as the nearer surface. Some degree of ambiguity as to the depth-difference is also indicated by the fact that 7 *O*s reported spontaneous reversals of the front-back relationship between the two surfaces at one time or another.

The amount of separation in depth between the two objects or the two surfaces was estimated on formal request by a sample of these *O*s, after the main procedure. The estimates were highly variable. For the spots they ranged from zero to 5 ft. For the textures they ranged from 2 in. to 3 ft. Some *O*s were unwilling to judge, saying, "It depends on what it is," or "It could be infinity." Evidently the impression of how far apart these entities were was indefinite, as also was the impression of how far away they were.

Figure 2 represents an ideal theoretical possibility of what the *O*s might have seen in this experiment, but it cannot be asserted that this is what they did see. The reports indicated that they perceived two things of some kind in some kind of space behind the screen, but neither the direction nor the amount of their separation in the third dimension was definite.

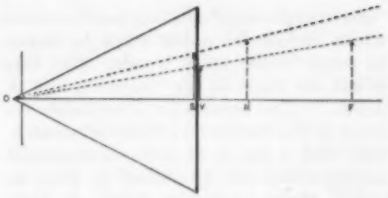

FIG. 2. The two angular velocities of Fig. 1, with a pair of virtual objects at different distances moving in apparent space. The two apparent objects (or surfaces) are shown as if seen behind the translucent window. The nearer (N) is shown close to the window; the farther (F) is shown at the distance it would have to be if the two were taken to be in rigid translation (if the two vectors were equal). In the diagram ON/OF = V/S, by trigonometry. That is, the distances are such that the optical velocities are inversely proportional to them. (Note that N corresponds to S, and F to V.) But this relation depends entirely on the assumption of rigidity. The absolute distances of N and F may vary, but the ratio of ON/OF will remain constant on that assumption. If rigidity is *not* assumed, however, there is no rationale for predicting any relation of distance or depth or even which object will be seen in front of the other.

Discussion

The significant result of this two-velocity experiment is not so much the effect of motion on depth perception as its effect in separating one surface into two. With the textures, all *O*s saw a single frontal surface when there was no differential motion, and all *O*s saw two frontal surfaces when there was a sufficient degree of differential motion. This separation is not what is ordinarily meant by depth, since it was not always clear which surface appeared in front and which behind.

The phenomenon is similar to the

"disentanglement" of foliage and branches which Helmholtz noted when he began to move in the thick woods. But this is not the same as his "apperception of depth." The separation is probably related to Wertheimer's (1923) demonstration that a group of spots interspersed among others will be unified by what he called their "common fate" if they moved together. Other conditions for the seeing of one thing in front of another, for transparency and superposition, have been discussed by Koffka (1935). Although the phenomenon may not seem relevant for some kinds of space perception it is certainly relevant for object perception.

How can this result be explained? Instead of appealing to a process of organization, or a law of "common fate," one might look for its basis in the geometry of the optical stimulus. Geometry distinguishes between (*a*) perspective transformation of forms, (*b*) topological transformations of forms, and (*c*) *disruptions*. These correspond roughly with the distinctions in physics between rigid motions of bodies, elastic motions of bodies, and the motions of breaking, tearing, or splitting. In this experiment, the motion of one set of textural elements relative to the other was a disruption, geometrically speaking. When sufficiently different velocities were imposed on them, the *adjacent order of elements* in the textures was destroyed. More exactly, there was a permutation of this order. It was a particular sort of permutation, to be sure, for each of two sets of elements retained an adjacent order, but the disruption of order as between these sets broke the original continuity. And this produced the perception of different surfaces with separation between. The detection by the eye of continuity or solidity as compared with discontinuity, disruption, or separation, is probably a fundamental kind of perception. The continuity of a single surface in two dimensions may be given by a static optical texture. But the continuity of a solid object in three dimensions probably depends on the kind of optical motion presented to the eye. Perhaps it was this lack of solid continuity or rigid connectedness between the nearer and farther surface in our experiment which prevented the ideal possibility represented in Fig. 2 from having been realized.

The earlier investigators of motion parallax were willing to assume that an eye was sensitive to the stimulus of motion, but they did not seem to realize that differential motion necessarily entails a *change of pattern*. In our experiment there were motions of the elements relative to the window but there were also motions of one set of elements relative to the other. For example, when both sets of elements were moving to the left, relative to the window, and one moved faster than the other, the slower was moving to the right, relative to the faster. Spontaneous reports of this appearance were given by several *Os*. Why should not a differential velocity be perceived just as directly as the two component velocities? When two moving elements are far apart in the field, one might suppose the slower and the faster velocity might have to be compared in order to detect a difference between them. But when the elements are adjacent in the field the difference is given by the change of pattern. Permutation of order is one type of change of pattern. In order to study the sensitivity of the eye to *form*, to *change* of form, and to the *forms* of change of form, a taxonomy of these variables is desirable.

Experiment II: Motion Parallax with a Flow of Velocities

Problem and Method

If a two-velocity field does not arouse consistent perceptions of depth, will a flow-velocity field do so? In order to make this comparison, the apparatus already described was modified so as to present to the eye a texture in which the horizontal velocities of the elements varied from slow to fast from the top to the bottom of the field, that is, a gradient. As before, the field could also be motionless.

Apparatus.—The shadows on the screen were produced by spattering paint on a transparent sheet interposed between the point

source and the screen. This mount was slanted toward the screen at an angle of 45° (Fig. 3). By the principle of mirror reversal, the gradient at the eye should yield an apparent slant *away* from the screen.

If this texture had been composed of elements regularly spaced in a grid or pattern, then the gradients of size and spacing would have been effective in producing the impression of a slanted surface even when the texture was motionless, as previous research has demonstrated (J. J. Gibson, 1955). But as it was, the brightness-transitions which composed the texture were not sharp, the elements had indefinite sizes and shapes, and the static gradients, if present, were not effective in producing a perception of slant. This texture was nevertheless sufficient to produce the perception of a continuous surface.

As before, *O* looked through an aperture which prevented his seeing either the edges of the window or any other part of the apparatus. The field was 82° wide by 52° high. The eye and the point source were each 75 cm. from the window.

Procedure.—All *O*s were naive. Each was told that when he applied his eye to the aperture he would see a gray field of view, and was asked to report what he saw. They were divided into two groups, one of which was presented with the moving texture without ever seeing it static (Group I), and the other with the moving texture after having seen it static (Group II). Their spontaneous descriptions were recorded and later classified. Questions were asked only after these reports, and in the same terminology used by *O*.

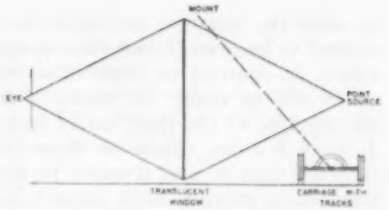

FIG. 3. The shadow projector giving a gradient of angular velocities, viewed from the side. The carriage and mount move back and forth parallel to the translucent window. The angular velocity is greatest at the bottom of the window, where the mount is closest to the point source, and least at the top, where the mount is farthest from the point source. At an inclination of 45°, this yielded a ratio of the minimum to the maximum velocity of 3 to 1. The apparent surface should appear to slant backward at the top.

Results

Nineteen out of 21 *O*s (Group I) reported a rigid moving plane surface of some kind slanting away from them at the top of the field. The inclination of the surface was estimated without difficulty, and they judged it to be receding upward. Of the 2 remaining *O*s, one's report was uninterpretable, and the other's was of a surface perpendicular to his line of sight.

With the static texture, 25 *O*s out of 28 (Group II) reported something which could be classed as surface-like, but which was in no case slanted backward into space. Of the remaining 3 *O*s, 2 saw a surface whose lower part was perpendicular but whose upper part was slanted back, and one saw the surface slanting back into space. When these 28 *O*s were later presented with the moving field, 26 saw it unambiguously to be a receding rigid surface at a fixed inclination. Of the remaining two, one saw it perpendicular to the line of sight; the other's report was uninterpretable.

Estimates of the slant of the apparent moving surface were obtained from each *O* in terms of degrees of inclination backward from the frontal plane. For Group I, the 19 judgments varied from 20° to 60° with a median of 40°. For Group II, the 26 judgments varied from 12½° to 55° with a median at 37½°. The theoretical value based on the gradient of velocities alone would be 45°. The medians show a constant error of underestimation. The surface never appeared to be at 90°, that is, it never looked like a ground on which *O* might be standing.

Estimates were requested of *how*

far away the apparent moving surface seemed to be from *O*, but these judgments, in contrast to those of slant, could not as easily be made. For the total of 49 *O*s, they varied from 3 in. to 5 miles. Some of these *O*s reported that it was possible to see themselves moving with respect to the surface instead of the surface moving with respect to them.

Discussion

These results indicate that motion parallax with a continuous gradient of velocities does induce consistent judgments of slant. This is the property of receding in depth in a certain direction. It is combined with the experience of a continuous rigid surface. The second experiment was like the first in most respects except that continuity of differential motion was introduced. To put it another way, permutation of textural elements was absent, but a skew of the pattern of elements was present. Judgments of distance in this experiment were highly variable and were made with reluctance, as they were in the first experiment.

There is evidently more than one kind of "depth" and more than one kind of motion parallax. The two experimental situations so far described were sufficient for judgments by naive *O*s of two kinds of spatial perception, (*a*) the phenomenon of separation in the third dimension and (*b*) the phenomenon of recession in the third dimension. Neither situation, however, was sufficient for absolute judgments of distance. This may be connected with the fact that no perception of a level ground was induced.

Experiment III: Correspondence Between Velocity-Pairs and Judgments of Depth Under Various Instructions

Problem and Method

In Exp. I, a difference in optical velocity did not induce the perception of a difference in visual distance with any great consistency. An explanation of this result has been suggested in terms of the absence of continuity between the two velocities. Another possible explanation however, in the spirit of Helmholtz, would be that the expected perception did not occur because *O* did not interpret the two motions as a difference in distance, or had not learned to perceive differential velocity as a difference in distance. On this theory, the suggestion that there was always an element of depth between the objects, with instructions about the limits of this separation, would be expected to alter the perceptions and lead to consistent judgments of depth. This prediction was tested in Exp. III, using the textures and spots of Exp. I.

Naive *O*s were given one of three degrees of suggestion or verbal information about depth, and were asked to reproduce on an adjustable but unmarked scale 1 m. in length the separation between the nearer and the farther apparent surface. The degree of velocity difference, the V/S ratio, was systematically varied so that judgments could be correlated with it.

Procedure.—All *O*s, as previously described, initially made spontaneous observations to the request "Describe what you see," in response to the motionless stimulus, the greatest velocity difference (V/S = .51), and the least velocity difference (V/S = .97). The apparatus was that of Exp. I. Each *O* then made 20 judgments of amount of separation for both the textures and the spots. The variable transparent mount was so set as to produce 10 velocity ratios of .51, .54, .57, .61, .65, .70, .76, .83, .90, and .97. Each ratio was presented twice, in random order, in the texture series and the spot series. Half the *O*s began with one series and half with the other. An *O* was assigned to one of three groups, and then instructed as follows:

Group I (*Least information*): The *O* was shown the sliding scale beside his chair and was told that it could be used to indicate the distance between the nearer and farther of the two (surfaces, spots). If *O* had used other terms instead of "nearer" and "farther," these terms were employed. He was then told that he would be shown a number of different settings of the apparatus, and that each time he would be asked to make a judgment (degree of separation, distance between, etc.). No other information was given. The *O* was encouraged to report as he went along but no comments were made on his performance. There were 16 *O*s.

Group II (*Maximum and minimum*): After the adjustable scale had been demon-

TABLE 1

MEDIANS AND RANGES OF CORRELATIONS BETWEEN VELOCITY RATIOS AND JUDGMENTS OF DEPTH

Measure	Stimuli	Group I (N = 16)	Group II (N = 17)	Group III (N = 20)
Median r	Surfaces	+.72	+.67	+.83
	Spots	+.57	+.51	+.84
Range	Surfaces	−.23 to +.95	+.34 to +.85	−.18 to +.96
	Spots	−.52 to +.94	−.14 to +.85	−.45 to +.99
Number r's significant at <.05	Surfaces	12/16	15/17	19/20
	Spots	14/16	8/17	19/20

strated, these *O*s were given another demonstration of the greatest and the least velocity-difference and were told which one was the maximum and which the minimum. The procedure thereafter was the same as for Group I. There were 17 *O*s.

Group III (*Most information*): After the preliminaries described, these *O*s were told: "These are shadows of two (surfaces, objects) which are actually at different distances from you. One is farther away from you than the other. This is what they look like at their maximum separation, which is 18 in. (The *O* was shown 18 in. on the adjustable scale.) This is what they look like at their minimum separation, which is ½ in. (The *O* was shown ½ in. on the scale.) Each time I will ask you to estimate how far apart the (surfaces, objects) making the shadows are." The procedure was thereafter the same as for the other groups. There were 20 *O*s.

When each *O* had finished his judgments, he was questioned by *E* as to how he had made them unless he had already made this clear. The questions were: How did you make your judgments? Did you go by appearance of depth, or did you try to use some other cue? If some other cue, what was it? Did you ever see the front and back (surfaces, spots) change places or fluctuate? Did the two (surfaces, spots) ever appear to be connected, like parts of a rigid object?

Results

For each *O* a rank order correlation was run between his 20 judgments of separation and the corresponding velocity-ratios. Table 1 summarizes the median coefficients for the three kinds of instruction and for the two kinds of apparent objects, surfaces and spots. They range from .51 to .84. Eighty-seven of the 106 correlations were significant at the 5% level ($r > .44$). The data thus demonstrate some correlation between amount of differential velocity and degree of depth judged. The spread of the *individual* correlations was very great, however, ranging from −.52 to +.99.

Group III, which was given the most information, had higher correlations (medians of .83 and .84) than the other two groups. If individual correlations are considered, 19 out of 20 were significant for both the surfaces and the spots. The correlations of Groups I and II do not differ from one another. The demonstrating of the maximum and minimum stimulus presentations at the ends of the scale did not, therefore, improve the ordering of the estimates. But telling *O* that he was seeing shadows of two things separated by a given amount of space did so.

When the instructions were given, they seemed to make *O* begin searching at once, and quite deliberately, for cues which he could put into some order. Only a minority of the *O*s reported that they had depended on any "appearance of depth" in making their estimates. Considering all *O*s, 70% reported that they had used

the "relative motion," or the "difference in speed," or "how far one passes the other" as the basis for judgment. Evidently, most of them saw motions of some kind of entities but did not clearly see the amount of space separating them.

The appearance of a connection between the two surfaces or spots "like parts of a rigid object" was almost never reported in answer to the question. The appearance of exchanging place or fluctuating in depth was reported by 17 out of 53 *O*s.

Because the more specific instructions raised the correlations, two *O*s (both psychologists and familiar with the apparatus and the problem) were run on consecutive days with reinforcement, to see how high the correlations might go and how stable they might become. The spots were not used in this experiment. The first 20 trials on each day were run without comment. But on the next 10 trials *O* was corrected after each judgment. The *E* did so by marking off on the scale the actual distance which separated the standard and variable mounts behind the translucent screen. This procedure of 20 uncorrected and 10 corrected trials was repeated for 9 days with one *O* and 7 days with the other.

The correlations improved rapidly from coefficients of .20 and .30 to ones of .90 or more. The introduction of new and unfamiliar textures had little effect on the correlation. Neither did the requiring of one *O* to use a verbal rating scale of 1 to 10 instead of the adjustable sliding scale.

These *O*s definitely took a problem-solving approach to the task and checked their methods of judging against the corrections given by *E*. One estimated the ratio of the two velocities and tried to express this numerically each time. The other said, after a few days, that the task had become similar to memorizing paried associates; he was trying to link up ordered "cues" with particular scale positions. The training did not have the effect of producing or enhancing an immediate appearance of depth.

Discussion

After verbal suggestion, information, or training concerning separation in depth, a correlation was present between the degree of velocity-difference and the degree of separation judged. It was raised by information and corrected training. But the reports indicated that the *O*s generally saw motions rather than depths, and that the appearance of depth was not induced by information or training. They were led to look for and use cues for the required judgments of depth, but not to report that they perceived it. They could interpret a velocity-difference as a depth-difference if given instructions, and they could improve the consistency of their estimates if given reinforcement. They could learn to perceive in one sense of the term, but they did not learn to see a differential velocity as a difference in depth.

This result does not support the theory of "unconscious inference" or point to any process for the conversion of bi-dimensional impressions into perceptions. It should be remembered, of course, that in these experiments motion has been isolated from other determinants of perception. This does not occur in everyday life. Motion usually comes in conjunction with size, shape, density, and disparity. But it might be supposed that the fact of this conjunction over years of past experience would have given an associative cue value to the motions in our experiment. In that case the velocity-pairs should have produced spontaneous judgments at once. Since they did not, this particular theory of cue learning is not supported by our results.

Summary

The common assertion that motion parallax is a cue for depth perception is vague. The optics of differential velocities of the elements in a field of view were examined and two cases were distinguished: that of *two* velocities in the field and that of a *gradient* of velocities in the filed. The two-velocity case yielded consistent perceptions of the separation of one surface into two. The flow-gradient case (motion perspective) yielded consistent perceptions of slant, or rate of recession in depth. In neither case were there consistent judgments of distance from *O*. Still another case, that of two different velocities of

two different spots in an otherwise empty field, did not yield consistent space perceptions of any kind. Evidently Helmholtz was wrong about an "immediate apperception" of the distances of bodies or the depth between them solely from an impression of the velocities of spots in the field of view. The only consistent or "immediate" impressions obtained were those of separation in depth (in response to a permutation of adjacent order of texture), and recession in depth (in response to a transformation of the texture).

These were spontaneous perceptions in the sense that no other information was given *O* than that carried in the optical stimulation. When he was given verbal information about depth in the two-velocity situation he was able to correlate a judgment of depth with an impression of motion. But the correlations were not perfect, and a minority of *O*s "saw" the depth. One might conclude from these facts that there are two kinds of optical stimulation for experiences of space: a kind which requires additional information to yield consistent judgments, and another kind which does not require it—which is compelling or coercive. The facts also suggest that there are two kinds of experience of space: "empty" depth, as exemplified by one surface in front of another, and "filled" depth, as exemplified by the slant or recession of a surface. The depth of a surface is perceived more consistently than is the depth of the space between surfaces, in our situation.

As regards the perception of absolute distances, this probably depends on the perception of a terrain or ground surface, the conditions for which were not reproduced in the present experiments. For the investigation of this problem, a very large field of view is required.

References

Bourdon, B. *La perception visuelle de l'espace.* Paris: Schleicher, 1906.

Gibson, J. J. The perception of visual surfaces. *Amer. J. Psychol.*, 1950, **63**, 367–384.

Gibson, J. J. Optical motions and transformations as stimuli for visual perception. *Psychol. Rev.*, 1957, **64**, 288–295.

Gibson, J. J., & Carel, W. Does motion perspective independently produce the impression of a receding surface? *J. exp. Psychol.*, 1952, **44**, 16–18.

Gibson, J. J., & Gibson, E. J. Continuous perspective transformations and the perception of rigid motion. *J. exp. Psychol.*, 1957, **54**, 129–138.

Gibson, J. J., Olum, P., & Rosenblatt, F. Parallax and perspective during aircraft landings. *Amer. J. Psychol.*, 1955, **68**, 372–385.

Graham, C. H. Visual perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951, 868–920.

Graham, C. H., Baker, K. E., Hecht, M., & Lloyd, V. V. Factors influencing thresholds for monocular movement parallax. *J. exp. Psychol.*, 1948, **38**, 205–223.

Helmholtz, H. *Physiological optics.* Vol. III. J. P. Southall, Ed., Optical Soc. America, 1925.

Hochberg, J., & Smith, O. W. Landing strip markings and the "expansion pattern": I. Program, preliminary analysis, and apparatus. *Percpt. mot. Skills*, 1955, **2**, 81–92.

Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.

Smith, O. W., & Smith, P. C. Interaction of the effects of cues involved in judgments of curvature. *Amer. J. Psychol.*, 1957, **70**, 361–375.

Tschermak-Seysenegg, A. Über Parallaktoskopie. *Pflüg. Arch. ges. Physiol.*, 1939, **241**, 454–469.

Vernon, M. D. Cognitive inference in perceptual activity. *Brit. J. Psychol.*, 1957, **48**, 35–47.

Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.*, 1923, **4**, 301–350.

(Received July 24, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

A COMPARISON OF THREE VARIETIES OF TRAINING IN HUMAN PROBLEM SOLVING[1]

LLOYD MORRISETT, JR.
Social Science Research Council
AND CARL I. HOVLAND
Yale University

Recently Adams (1954) compared the efficiency of single-problem training and multiple-problem training in producing transfer to a new problem. He showed that single-problem training resulted in more transfer than did the multiple training procedure. Approaching the same issue, but using different experimental techniques, Callantine and Warren (1955) have shown that the greater the number of training problems the greater will be the amount of transfer on a test problem. These results seem contradictory.

Apart from differences in the problems used, apparatus, and *S*s, an essential difference in the two experiments appears to be in the degree of mastery of the training task before the test for transfer. As Callantine and Warren point out, all of Adams' single-problem training *S*s mastered the training task while many of his multiple-problem training *S*s did not. Their *S*s, on the other hand, were trained to a criterion so that on any trial they continued until they made the correct response.

This difference in procedure is analogous to the differences in procedure between the experiments of Harlow and Warren (1952) and Hayes, Thompson, and Hayes (1953). Analysis of these superficially discrepant results shows that two factors seem to be necessary to produce a high degree of transfer in these "learning set" types of situation: (*a*) a moderate to high degree of learning must be achieved on a single problem (Adams, 1954; Hayes, Thompson, & Hayes, 1953), and (*b*) there must be presentation of several problems in order to insure learned generalization (Callantine & Warren, 1955; Harlow & Warren, 1952; Hayes, Thompson, & Hayes, 1953). In addition there is a third factor to consider. Research by Wolfle (1935; 1936) and Grether and Wolfle (1936) has also shown that increasing the amount of variation in the stimulus from trial to trial decreases the rate of learning. These factors suggest that the reason Adams' multiple-problem training group was inferior to his single-problem training group was the low degree of learning achieved by the multiple group within any *one* training problem; an experimental procedure involving only two trials per problem would be expected to produce little learning on any one problem.

From this analysis, taking all three factors into account, it is first hypothesized that *S*s who have a chance to achieve a high degree of learning within a single problem and also receive training on more than one problem will be superior to both a single-problem training group and a multiple-problem training group of the type used by Adams. Extra-

[1] This study was carried out while the first author was a N. S. F. Predoctoral Fellow in psychology at Yale University studying under the direction of Carl I. Hovland.

polating from Adams' results, a second minor hypothesis is also formed: other things being equal, single-problem training will result in more transfer of training than will multiple-problem training when learning within any one problem in the multiple training procedure is minimal.

METHOD

In order to test the two hypotheses three groups of *Ss* were compared in an experiment which in procedure essentially duplicated that of Adams (1954). All *Ss* received the same number of training trials, 192. The three groups of *Ss* differed in the distribution of training trials. Group I received all of its training, 192 trials, on a single problem. Group II was trained for 8 trials on each of 24 different problems. Group III was trained for 64 trials on each of 3 problems. The *Ss* received 24 test trials following training. The test problem was the same for all *Ss*.

As in Adams' experiment, training and testing consisted in presenting each *S* with a series of easily discriminable pairs of geometrical figures,[2] e.g., a circle and a triangle. The *S* was required to choose a response, the correctness of which depended on the spatial arrangement of the stimulus pairs. The stimuli were presented in four horizontal-vertical arrangements, and *S* had the opportunity to make any one of four responses to each pair. For example, the circle could be either to the left, right, above, or below the triangle. For each of these arrangements a different response was correct. The *Ss* responded by pressing one of four numbered buttons set in a key box[3] before them. If the correct response was made, a small bulb by the key box lighted. If the response was incorrect, nothing happened.

A single pair of geometrical figures constituted a problem. Since there were four horizontal-vertical arrangements of a single pair, the complete presentation of a problem required four trials. Group I had, then, 48 complete presentations of its single problem in training, while Group II had two complete presentations of each of 24 problems in training, and Group III had 16 complete presentations of each of its three training problems.

The stimulus pairs were presented on slides at a rate of one slide each 10 sec. using a Revere automatic projector. Within blocks of eight the slides were randomized with the restriction that no adjacent slides be identical. When *S* pressed the correct button for a stimulus pair his response was recorded on an Esterline-Angus recorder. The use of slides, the key boxes and lights, and a 20-channel Esterline-Angus recorder made it possible to run up to 20 *Ss* at one time.

The *Ss* were obtained from the twelfth-grade class of the East Haven, Connecticut, High School.[4] They were assigned randomly to groups for the experiment. While originally equal numbers of *Ss* had been assigned to each group, apparatus failures and school absences resulted in the present unequal distribution of *Ss*. Group I contained 19 *Ss*, Group II contained 12 *Ss*, and Group III contained 32 *Ss*.

When *Ss* arrived for the experiment they were each seated in front of a key box. The procedure of showing pairs of geometrical figures and the operation of the key box and lamp were briefly explained. The *Ss* were cautioned to press only one button for each slide and were told that all responses were being recorded.[5] After the instructions were read and questions answered, the experiment proceeded. The slides were shown without interruption except for brief pauses of 10 sec. following Slides 36, 72, 108, 144, and 192 for changing the magazine in the projector.

RESULTS

Figure 1 shows the mean errors for each of the three experimental groups during training and in the test series; the errors are shown for blocks of eight trials. Since the distributions of errors are skewed, the statistical evaluation of the data has been made using appropriate nonparametric methods.

[2] The authors wish to thank Jack Adams for making it possible for them to use the same geometrical figures which he had used.

[3] The authors wish to express appreciation to the Southern New England Telephone Company for the use of their key boxes in this experiment.

[4] Carl Garvin, Principal of East Haven High School, generously cooperated with the authors in making this study possible.

[5] Actually, *Es* had no way to prevent the pressing of more than one button per slide or to record this kind of behavior. However, even though this practice probably occurred it would appear not to be a differential factor among the various groups.

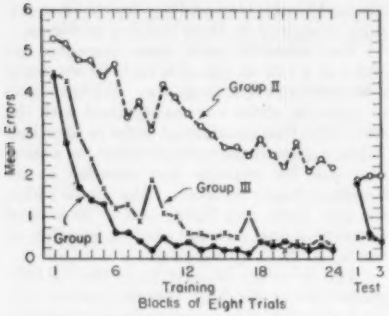

FIG. 1. Mean errors, by blocks of eight, on training and test trials.

Examination of Fig. 1 shows that the error scores of all three groups diminished during training, with, as expected, Groups I and III improving more rapidly than Group II. The temporary decrements in the performance of Group III on Blocks 9 and 17 coincide with the introduction of the second and third training problems for that group. Kruskal and Wallis' (1952) *H* statistic was used to test for possible initial differences among the groups at the start of training. These differences were not found to be significant ($P > .30$).

The two original hypotheses suggest two critical comparisons as tests. The major hypothesis was that training which gives a chance for both learning within a problem and generalization between problems will be superior in producing transfer to training which allows only one of these mechanisms to operate; Group III should make fewer errors in the test series than either Groups I or II. The second hypothesis was that, other things being equal, more transfer will result from training which produces a high degree of learning within a single problem than from training on many problems with only a minimal amount of learning on each; Group I should make fewer errors in the test series than Group II. Whitney's (1951; Mosteller & Bush, 1954) extension of the *U* statistic allows a simultaneous test of these two hypotheses within the framework of the over-all design. This test rejects the null hypothesis, III = I = II, in favor of the predicted alternative, III < I < II ($P < .01$). Both original hypotheses are therefore supported.

With reference to Fig. 1, it will be seen that the superiority of Group III over Group I is almost entirely due to errors on the first block of the test series. The single-problem training group, I, is not initially superior to the multiple-training group, II, but the performance of Group I improves rapidly in the series of test trials and by Block 3 is equivalent to that of Group III. This is similar to the result obtained by Adams. In that experiment he found his single-problem training group not to be initially superior to his multiple group in the transfer test, but in the later stages of the transfer test the single-problem group improved more rapidly than the multiple-problem group and was significantly superior.

DISCUSSION

The present results are seen as reconciling the apparent discrepancy between the experiments of Adams (1954) and Callantine and Warren (1955). When the noncorrection technique used by Adams is employed there is likely to be a low degree of learning within any block of trials, and the learning of a large number of problems at a low level of mastery is less efficient in producing transfer than concentration on thorough learning of a single problem. Hence, continued practice is essential to produce adequate original learning. On the other hand, the correction technique used by Cal-

lantine and Warren insures a high degree of learning within each of the multiple problems, and here the greater the variation in stimulus conditions the greater the transfer.

Results involving transfer after training on a number of successive tasks are typically explained in terms of "learning to learn" (Harlow, 1949; McGeoch & Irion, 1952). This seems to imply that new habits are set up within each problem. We are more inclined to conceptualize the process as the learning of a single class of habits to a complex set of cues. Successful transfer would involve learning responses to the complex pattern of stimuli produced both by the immediate environmental stimuli, e.g., the pattern of the slide, and traces from prior stimuli. Specifically, in the present type of problem the following must be accomplished: (*a*) one must learn that up-down variation is associated with Responses 1 and 2, while left-right variation is associated with Responses 3 and 4; (*b*) one must learn that one figure of the stimulus pair, e.g., the circle, is associated with Responses 1 and 3, while the alternate figure, e.g., the triangle, is associated with Responses 2 and 4; and finally (*c*) proficiency must be acquired in making the appropriate response to the combined stimulus. Thus, traces of prior stimuli as well as immediate environmental stimuli determine the response.

With this formulation it is clear why both a high degree of learning on each single problem and stimulus variation are necessary. A high degree of learning within a single problem provides for a strong stimulus-response association. If prior stimulation is to guide subsequent responses, however, the relevant stimulus traces must be discriminated from irrelevant stimuli. This means training on a second problem where the stimulus objects differ, and differential reinforcement can produce discrimination. After extensive variation in the stimulus materials the stimulus trace from a correct choice, or an incorrect choice, will be a relatively invariant, discriminated part of the stimulus complex.

Summary

Sixty-three *Ss*, divided into three groups, were given different distributions of 192 discrimination problem training trials and then tested on the same transfer problem. Groups I, II, and III were trained on 1, 24, and 3 problems, respectively. On the transfer test Group III proved to be superior to both Groups I and II, while Group I was superior to Group II.

References

Adams, J. A. Multiple versus single problem training in human problem solving. *J. exp. Psychol.*, 1954, **48**, 15–19.

Callantine, M. F., & Warren, J. M. Learning sets in human concept formation. *Psychol. Rep.*, 1955, **1**, 363–367.

Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.*, 1952, **47**, 583–621.

Grether, W. F., & Wolfle, D. L. The relative efficiency of constant and varied stimulation during learning. II. White rats on a brightness discrimination problem. *J. comp. Psychol.*, 1936, **22**, 365–374.

Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.

Harlow, H. F., & Warren, J. M. Formation and transfer of discrimination learning sets. *J. comp. physio. Psychol.*, 1952, **45**, 482–489.

Hayes, K. J., Thompson, R., & Hayes, C. Discrimination learning sets in chimpanzees. *J. comp. physio. Psychol.*, 1953, **46**, 99–104.

McGeoch, J. A., & Irion, A. I. *The psychology of human learning.* New York: Longmans, Green, 1952.

Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology.* Cambridge, Mass.: Addison-Wesley, 1954. Pp. 289–334.

Whitney, D. R. A bivariate extension of the *U* statistic. *Ann. Math. statist.*, 1951, **22**, 274–282.

Wolfle, D. L. The relative efficiency of constant and varied stimulation during learning. *J. comp. Psychol.*, 1935, **19**, 5–27.

Wolfle, D. L. The relative efficiency of constant and varied stimulation during learning. III. The objective extent of stimulus variation. *J. comp. Psychol.*, 1936, **22**, 375–381.

(Received July 25, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

REVERSAL AND NONREVERSAL SHIFTS IN KINDERGARTEN CHILDREN[1]

TRACY S. KENDLER
Barnard College
AND HOWARD H. KENDLER
New York University

The purpose of this study was to determine whether kindergarten children would respond to the experimental task in a manner consistent with a single unit or a mediational S–R theory of discrimination learning. The experimental task involved learning an initial discrimination between stimuli that differed on two dimensions. Only one dimension was relevant, i.e., rewarded. After the criterion was reached a second discrimination was presented. For reversal shift *Ss* the same dimension remained relevant but the reward patterns were reversed. For the nonreversal shift *Ss* a previously irrelevant dimension became relevant.

The single-unit S–R theory is a name applied to formulations that, like Spence's (1936) discrimination learning theory, assume a direct association between physical stimuli and overt responses. This theory predicts that requiring an organism to make a response opposite to that which it has previously learned, as in a reversal shift, should result in slower learning than a nonreversal shift. Kelleher (1956), using albino rats as *Ss*, has corroborated this prediction.

Spence has been quite explicit about limiting his predictions to inarticulate organisms. Human beings learn to make verbal or symbolic responses, overt or covert, to physical stimuli. These responses produce stimuli that mediate between the external stimulus and the overt response. Kendler and D'Amato (1955) applied a mediational theory to predict the results of reversal and nonreversal shifts in a concept formation study on college students, using a card sorting technique. They reasoned that "a reversal shift should occur at a more rapid rate than a nonreversal shift because at the completion of the learning of the first concept, the symbolic cues appropriate to the second concept were available to *Ss* in the reversal shift groups" (1955, p. 169). Their prediction was confirmed. Thus, if a card sorting situation can be compared with a discrimination learning situation, it appears that reversal shift results in positive transfer for human adults and negative transfer for rats.

In the present research Kelleher's experimental procedure and design have been adapted for use with children to determine whether it would be possible to identify a transitional process in the development of concept formation in children from behavior rendered accountable by a single-unit S–R theory to the more complex responses associated with the mediational-sequential S–R theories.

METHOD

Subjects.—The *Ss* were 122 kindergarten children (age range from 58 to 78 mo.) from Willets Road Public School, Nassau County, New York. Fifty-three of the children were

[1] This research was sponspored by a grant from the National Science Foundation.

girls and 69 were boys. Two of the girls were eliminated because of their failure to learn the first discrimination. All *Ss* were randomly distributed to the various groups without regard to age and sex.

Apparatus.—The apparatus was an unpainted turntable made of a board 12 × 16 in., mounted on a swivel base and divided in half by a perpendicular plywood board 8 in. high and 15 in. wide. On one of the halves were two felt-padded depressions 1 in. sq. and 4 in. apart. The marble that served as the token reward was placed on one of the pads and covered by the opaque discriminanda. While these manipulations were in process the felt-padded half of the turntable was toward *E*. The perpendicular board served to screen *E*'s actions. When the stimuli were correctly arranged a swinging screen was brought down, concealing the discriminanda as they were swivelled into place in front of *S*.

The discriminanda for the initial training of the experimental groups and the test training for all groups were four metal tumblers 2½ in. in diameter that varied both in height and brightness. Two of the tumblers were 6¼ in. high (T). The other two were 4¾ in. high (S). One of each size was enamelled black (B) and the other white (W). The discriminanda for the initial training of the control group were four gray enamelled metal cookie cutters, two diamond and two circle shaped. One of each shape had two round, colored paper dots pasted on its upper surface.

The final items of apparatus were two marble boards, each with 50 holes capable of holding the marbles won by the child.

Procedure.—Children were run individually in a room in which *E* and *S* were alone. The *E* sat facing *S* with the apparatus on a small table between them. When the child was comfortably seated *E* said, "This is the game we are going to play, (child's name). Before we start, listen carefully and I will tell you how the game is played. See, there are two things here. When we start the game you will choose one of them and pick it up. If you are right you will find a marble under it. If you are wrong then you won't find anything under it. Each time you may choose only one. Then I will close it, turn it around like this and then you will have another turn. But on each turn you may choose only one. The game is to see how soon you can get a marble every time you choose. If you get the marble you put it into one of these holes. When we finish the game you may choose one of the prizes over there." The *E* allowed time for *S* to survey the assortment of prizes, that consisted of two chocolate bars (one with almonds and one without), a bag of salted peanuts, a box of raisins, a box of animal crackers, and 10 gold charms. Then *E* said, "Remember, the game is to see how soon you can get a marble every time you choose." This sentence was repeated from time to time during the course of the experiment. All but eight *Ss* completed the experiment in one session. Of these eight *Ss* who required two sessions, four were nonreversal *Ss*, three were reversal *Ss*, and one was a control *S*.

For the training discrimination, *Ss* were randomly assigned to one of three groups. The 40 *Ss* of the control group were presented with the shape discrimination. For half, the diamond was positive, for the others the circle was positive. The presence or absence of the dots was an irrelevant cue. The experimental *Ss* were each presented with the tumblers paired as follows: SB-TW and TB-SW. Forty *Ss* learned to respond to brightness (B vs. W). For half, B was positive, for the other half W was positive. The remaining 40 *Ss* learned to respond to height, half to T and half to S.

The test discrimination was presented immediately after the criterion was reached, with no change in instructions or interruption in the procedure. Each of the experimental groups was divided equally into a reversal (R) and nonreversal (NR) group, equated for initial speed of learning. The same tumblers used in the initial discrimination were used as stimuli, but they were now paired so that on any one presentation they differed on only *one* stimulus dimension.[2] The R *Ss* whose initial training was on the brightness dimension and NR *Ss* whose initial training was on the height dimension were presented with TB-TW and SB-SW successively. For the R group, if W was previously correct B was now rewarded, and vice versa. For the NR group, half were rewarded for B and half for W choices.

Similarly, the R *Ss* for whom height was initially relevant and NR *Ss* for whom brightness was relevant were now presented with SW-TW and SB-TB. For the R group, if S

[2] Buss (1953) has pointed out that if the stimuli in the second discrimination are paired as they were in the first, the nonreversal shift *Ss* are intermittently reinforced for the previously correct response. This change in the manner of presentation, which was adapted from Kelleher (1956), serves to minimize the retarding effect of such inadvertent reinforcement of a response that is no longer correct.

TABLE 1

MEAN NUMBER OF TRIALS TO CRITERION ON TEST DISCRIMINATION FOR EACH DIMENSION

Group	Dimension		
	Size	Brightness	Combined
Reversal	13.5	16.8	15.2
Nonreversal	7.5	17.0	12.2
Control	6.2	7.9	7.1

was previously correct T was now rewarded, and vice versa. For the NR group half were rewarded for T and half for S choices. In short, the reversal *Ss* were required to respond to a stimulus feature which had been consistently nonreinforced. The nonreversal *Ss* were required to respond to a stimulus feature that had received approximately the same number of reinforcements as nonreinforcements.

As for the control group, half was presented with TB-TW and SB-SW, with one-quarter rewarded for B and the other for S. The other half of the control group was presented with SW-TW and SB-TB, with one-quarter rewarded for T and the other for S.

Each cue appeared an equal number of times on the right and left. No combination of two cues appeared together on more than two consecutive trials. To reduce position effects during initial training, the position of the stimuli remained unchanged from trial to trial until a correct response was made. This practice was maintained only during the training and *not* during the test discrimination. The criterion of learning was 9 out of 10 successive correct responses.

RESULTS AND DISCUSSION

The response measure used was the number of trials required to reach the criterion. The means for the training discrimination for the R and NR groups were 18.77 and 17.77, respectively. This difference was not statistically significant.

Table 1 reports the data of the test discrimination. Five *Ss*, three in the R group and two in the NR group, did not meet the criterion. Their training was discontinued after 100 trials, and they were arbitrarily assigned a score of 100. Two of these *Ss*, one R and one NR, were among those who required two sessions.

Because the distributions were skewed and the variances heterogeneous, a $\sqrt{X + .5}$ transformation was applied to the raw data to prepare them for a factorial analysis of variance. Neither of the main effects (the type of transfer or the dimensions transferred to), nor the interaction between them was statistically significant.

The theory that is based on the assumption of a direct connection between the physical stimulus and the overt response (the single-unit theory) predicts that reversal should be slower than nonreversal or control. The mediational theory predicts that reversal should be faster than nonreversal. Our results confirm neither prediction. One conclusion we can draw is that neither theory is applicable, but another more interesting possibility is that both apply. The failure to achieve results consistent with either those obtained from rats (Kelleher, 1956) or from college students (Kendler & D'Amato, 1955) may be due to the fact that these children are in a transitional stage of development, in which these tasks lead some to function on a single-unit S-R basis, while others are making relevant mediated responses. If these two groups were about evenly divided, they would yield total results such as we obtained.

No MA scores were available. But there was a measure of performance that was perhaps more relevant to the task in hand; namely, the number of trials to criterion in the training discrimination. The *Ss* were sorted into two groups on the basis of their initial performance on each dimension. Those scoring above the median, i.e., taking more trials to learn, were labelled slow learners. The *Ss* scoring below the median were labelled fast learners. The results of fast and slow learners were

analyzed separately and are presented in Table 2.[3]

The difference between fast and slow learners could be accounted for along two possible lines. One of these is that the speed of initial learning reflects differences in general intelligence or ability to learn. If this were the case, it would be expected that results of the test discrimination for the fast and slow learners of the control group would show a difference that reflected the ability displayed in the initial discrimination. The obtained difference is both negligible and statistically unreliable, casting some doubt on the general learning ability explanation.

Another possibility is that fast learners approached the experimental task with verbal labels for the correct stimulus already strongly attached, due to prior experience with these stimuli. Liublinskaya (1957) demonstrated that attaching a verbal label to the distinguishing feature of the discrimination greatly facilitates learning. The slow learners in the present experiment may not have had appropriate verbal labels as strongly attached to the relevant stimulus as did the fast learners.

Regardless of whether the difference between fast and slow learners is based on general learning ability or the availability of relevant mediating responses

TABLE 2

MEAN NUMBER OF TRIALS TO CRITERION ON TEST DISCRIMINATION FOR *Ss* SCORING ABOVE AND BELOW THE MEDIAN ON THE TRAINING DISCRIMINATION

Group	Performance on Training Discrimination	
	Above Median (Slow Learners)	Below Median (Fast Learners)
Reversal	24.4	6.0
Nonreversal	9.0	15.5
Control	7.3	6.8
Combined	13.6	9.4

[3] The same analysis using errors instead of trials to criterion produced essentially the same result.

TABLE 3

ANALYSIS OF VARIANCE OF TEST DISCRIMINATION SCORES TRANSFORMED TO $\sqrt{X} + .5$

Source	*df*	*MS*	*F*
Type of transfer (R, NR, C)	2	12.17	2.14
Initial learning (fast, slow)	1	10.00	1.76
Interaction	2	22.74	4.00*
Within cells	114	5.69	

* $P = < .05$.

(which may in the long run prove to be synonymous with learning ability), the expectation was that fast learners would, like the college students, perform according to the mediational S-R theory, and therefore reversal would be faster than nonreversal. Slow learners were expected to behave more like Kelleher's rats, in accordance with the single-unit S-R theory, and therefore reversal would be slower than nonreversal. Table 2 shows the confirmation of these expectations. The analysis of variance of the transformed scores,[4] as presented in Table 3, suggests that the predicted interaction is statistically reliable ($P < .05$). Statistical subdivision of the interaction term (Walker & Lev, 1953) revealed that the variation in this term arose almost exclusively from the shift in the relative efficiency of the reversal and nonreversal conditions for the fast and slow learners.

On the basis of the results of the total groups and the analysis into fast and slow learners, it seems reasonable to conclude that these children, taken as a group, were in the process of developing mediating responses relevant to this task and that some were further along than others.

SUMMARY

The experiment compared the performance of kindergarten children in reversal and

[4] A similar analysis of variance was performed on the raw scores to test for possible distortions due to the transformation. The raw scores yielded almost identical results.

nonreversal shifts in a discrimination learning situation. The results obtained for the total group failed to support either a single-unit or mediational S-R theory, since there was no significant difference between reversal and nonreversal shifts. When the group was divided into fast and slow learners on the basis of their performance on the initial training series, suggestive differences were revealed. Fast initial learners responded in a manner consistent with a mediational S-R theory, i.e., reversal was faster than nonreversal. Slow initial learners responded in a manner consistent with a single-unit S-R theory, i.e., reversal was slower than nonreversal.

References

Buss, A. H. Rigidity as a function of reversal and nonreversal shifts in the learning of successive discriminations. *J. exp. Psychol.*, 1953, **45**, 75–81.

Kelleher, R. T. Discrimination learning as a function of reversal and nonreversal shifts. *J. exp. Psychol.*, 1956, **51**, 379–384.

Kendler, H. H., & D'Amato, M. F. A comparison of reversal shifts and nonreversal shifts in human concept formation behavior. *J. exp. Psychol.*, 1955, **49**, 165–174.

Liublinskaya, A. A. Development of children's speech and thought. In B. Simon (Ed.), *Psychology in the Soviet Union*. Stanford: Stanford Univer. Press, 1957. Pp. 197–204.

Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.

Walker, H. M., & Lev, J. *Statistical inference*. New York: Holt, 1953.

(Received July 30, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

EFFECT OF MAGNITUDE OF REINFORCEMENT ON ACQUISITION AND EXTINCTION OF A RUNNING RESPONSE [1]

HARVARD L. ARMUS

Institute of Psychiatric Research, Indiana University Medical Center, Indianapolis

Zeaman (1949) found starting time in a runway to be inversely related to magnitude of food reinforcement during both acquisition and extinction. Metzger, Cotton, and Lewis (1957) and Crespi (1942) found similar relationships for acquisition using total response time. The present study was designed to provide further data on these relationships, with total response time broken into starting time and running time.

METHOD

Subjects and apparatus.—The *S*s were 28 male Sprague-Dawley rats, aged 128–136 days, divided into two groups of 13 and 15. The apparatus was an enclosed, 3-ft. runway with a 9-in. starting box and an 18-in. right-angled goal box. The entire apparatus, with the exception of the unpainted hardware cloth lids, was painted dark grey. Time measurements were taken photoelectrically. The light beams were located immediately beyond the starting box door and immediately before the goal box door. Doors were of the guillotine type.

Procedure.—After approximately 2½ weeks on a feeding schedule of 1½ hr. on dry Purina Laboratory Chow checkers every 24 hr., each *S* was given two daily 10-min. exploration sessions, food and food cup being absent from the apparatus. During training and extinction each *S* was given five daily trials, spaced approximately 15 min. apart, while under 20–21-hr. food deprivation. Group 1 was reinforced with one 45-mg. food pellet; Group 10, with 10. On each trial, *S* was placed into the starting box facing the door. After 5 sec., the starting box door was opened regardless of the position of *S*. There were 15 days of acquisition and 30 of extinction.

[1] The author wishes to thank Elizabeth Ives for collection and computation of the data and James Norton for the design of the statistical analysis.

RESULTS

As the greatest differences would be expected during the later stages of acquisition and the early stages of extinction, the statistical analysis deals only with the last one-third of the acquisition period and the first one-third of the extinction period, except as otherwise noted. The complete data are presented in Fig. 1 and 2. Because of the presence of extreme scores, resulting in marked skewness and heterogeneity of variance, all probability values were determined by means of the Mann-Whitney *U* test.

Group 10 was superior to Group 1 during the relevant acquisition period in terms of both starting time ($P < .001$) and running time ($P < .001$). For extinction running time the reverse relationship held, Group 10 showing poorer resistance to extinction than Group 1 ($P = .03$). Extinction starting time yielded no significant differences ($P > .10$).[2] The interaction between amount of reinforcement and condition (acquisition-extinction) was significant for both starting time ($P < .02$) and running time ($P < .02$).[3] For neither re-

[2] Frequently during extinction a rat would dart partly out of the starting box, breaking the first beam, and then return to the starting box for a period of time. Even on such trials, the breaking of the first beam defined the end of starting time and the beginning of running time.

[3] The interactions were computed as follows: Each rat was assigned a difference score consisting of its mean time (starting or running) for the first third of extinction minus

sponse measure were the group differences for the first acquisition trial or first acquisition day significant. The running time data for the first extinction day (Trials 2–5) revealed nonsignificantly greater extinction for Group 10 than for Group 1, while a nonsignificant difference in the opposite direction was found for the second extinction trial of this day.

Discussion

The data reveal that, at least for the somewhat less variable running time measure, the relation between amount of reinforcement and performance is reversed from acquisition to extinction. Although Zeaman (1949) reported starting times only, he stated that his unreported running time findings were similar to, but not as significant as, the starting time data. The results of the present study agree with previously reported acquisition data (Crespi, 1942; Metzger, Cotton & Lewis, 1957; Zeaman, 1949), but do not support Zeaman's extinction findings.

There is at least one difference in procedure between the present study and that of Zeaman which might be of some importance in accounting for the discrepancies in the results. Zeaman's animals received 18 trials with the appropriate amount of reinforcement, as contrasted to 75 in the present study. It is likely that the effects of magnitude of reinforcement on extinction are partly determined by the number of exposures to the reinforcement magnitude before the start of extinction. However, it is problematic whether manipulation of this variable would alter the *relationships* during extinction between groups having

its mean time for the last third of acquisition. These difference scores were then ranked and the *U* test applied.

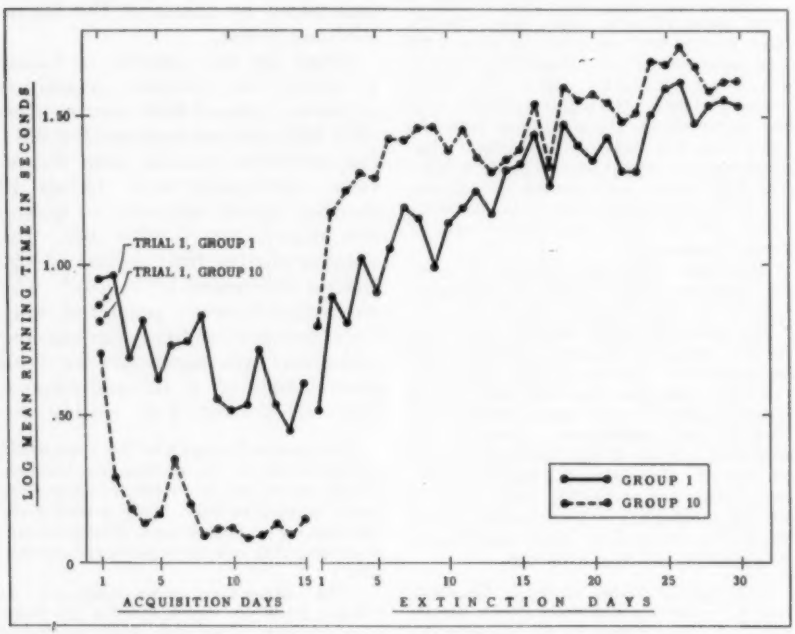

Fig. 1. Log mean running time during acquisition and extinction.

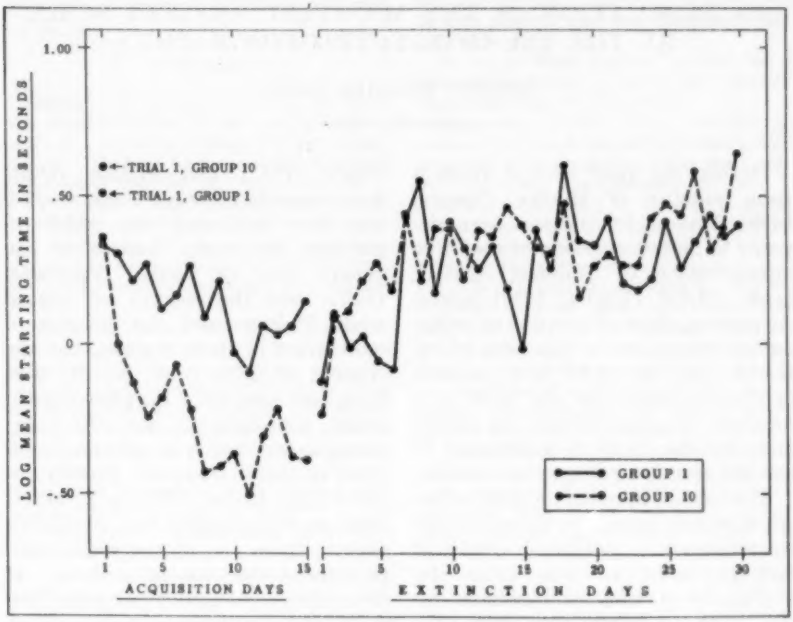

FIG. 2. Log mean starting time during acquisition and extinction.

had different magnitudes of reward during acquisition.

SUMMARY

In rats given runway training under 20–21-hr. food deprivation, it was found that the group reinforced with ten 45-mg. food pellets showed faster acquisition in terms of both starting and running times than the group reinforced with only one such pellet. During extinction, the reverse relationship held for running times; starting time differences were not significant. While the acquisition results supported previous findings, the extinction data did not.

REFERENCES

CRESPI, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.

METZGER, R., COTTON, J. W., & LEWIS, D. J. Effect of reinforcement magnitude and order of presentation of different magnitudes on runway behavior. *J. comp. physiol. Psychol.*, 1957, **50**, 184–188.

ZEAMAN, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received July 31, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

AN EVALUATION OF THE MANIFEST ANXIETY SCALE BY THE USE OF ELECTROMYOGRAPHY [1]

ASCANIO MICHAEL ROSSI [2]

University of Utah

During the past several years a large number of studies (Spence, 1958; Taylor, 1956) have been devoted to an investigation of the drive interpretation of Manifest Anxiety Scale (MAS) (Taylor, 1953) scores. At present, three alternative hypotheses are dominant in this area of research: (*a*) the MAS is a measure of chronic drive; (*b*) the MAS is a measure of susceptibility to drive; and, (*c*) the MAS is a measure of learned reaction patterns to anxiety.

The conflict between these hypotheses does not appear to be resolvable by reference to published empirical findings, since in these cases the evaluation of what is measured by the MAS is confounded with the evaluation of some hypothesis derived from a theory. The resolution of this conflict will probably be accomplished only when the evaluation of what is measured by the MAS is carried out independently of any test of a theoretical hypothesis. Jessor and Hammond (1957) have elaborated more fully on this point.

The results of another group of studies, which investigated the relationship between muscle action potentials (MAP) and performance, suggest a means by which this independent evaluation may be carried out. Duffy (1957) and Malmo (1957) have recently reviewed these studies, and they concluded that MAP recordings are useful indices of the energy level of human organisms. Duffy used the concept of arousal while Malmo used the concept of motivation in these reviews, but the concept of drive could equally well have been used since the phenomenon under consideration was the energizing aspect which is common to all three of these concepts. Several authors (e.g., Farber, 1954, p. 2; Marx, 1956, p. 97) consider this energizing aspect to be one of the main components of the concept of drive. If drive energizes reaction patterns, then it is logical to assume that it must first activate the organism's means for reacting, i.e., the skeleto-muscular system. Of course, there are several ways in which an organism may react without an involvement of the musculature (e.g., increase in brain activity). However, most psychological theorists who have utilized the concept of drive have been primarily interested in behavior which involved the skeleto-muscular system. The theoretical and empirical relationships between MAP and the concept of drive, discussed more fully elsewhere (Rossi, 1958), indicate that MAP recordings would be valid criteria for measuring drive levels. Accordingly, this measurement was used in the present study as a criterion measure in evaluating the following hypotheses: (*a*) the MAS is a measure of chronic drive; (*b*) the MAS is a measure of susceptibility to drive; and, (*c*) the MAS is a

[1] This study is part of a dissertation submitted to the faculty of the University of Utah in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to express his sincere appreciation to L. E. Bourne, Jr. for his valuable advice and criticisms during all phases of this study.

[2] Now at Harvard Medical School, Department of Psychiatry.

measure of learned reaction patterns to drive.

Method

Subjects.—A modified form of the MAS was administered to 179 men and 140 women students enrolled in introductory psychology classes. The form consisted of the 50 critical MAS items plus the 15 items from the L ("Lie") scale in the Minnesota Multiphasic Personality Inventory (MMPI). Correlation coefficients of .85 and .95 have been reported between a similarly modified form of the MAS and the full 225 item form (Shore, 1958).

Thirty of the women students were employed in a pilot study, and the experiment proper included 36 volunteer men. Eighteen of the men had achieved a score of 20 or above on the MAS and were designated "HI-MAS." The remaining 18 men had achieved a score of 8 or less on the MAS and were designated "LO-MAS." The only other restriction imposed on the selection of *Ss* was that no *S* had obtained an L score above 7.

Tasks.—Two reaction time (RT) tasks were devised (simple and complex), and each of the tasks was administered under two conditions (threat and no threat). The nature of the tasks and conditions prohibited a counterbalancing of the order of presentation. However, a partial counterbalancing effect was achieved by administering one order of tasks to one-half of the *Ss* in each MAS group, and another order of tasks to the other half. Since the tasks were administered in close temporal proximity the threat and no threat condition could not be effectively counterbalanced. The no threat condition always preceded the threat condition.

The *simple task* required *S* to push down his left switch when the left stimulus light came on, and to push down his right switch when the right stimulus light came on. The *complex task* required that *S* push down the right switch when the left stimulus light came on, and to push down the left switch when the right stimulus light came on.

In the *threat condition* an electrode, in addition to those used for MAP recording, was attached to *S*, who was instructed that the switches had been adjusted in accordance with the average RT for the girls in his class. He was told that if he reacted slower than "the average girl in his class," he would receive a shock, and if he reacted faster than "the average girl in his class," he would escape the shock. In fact, however, no shock was administered. The instructions probably had the double effect of threatening *Ss* with physical pain and threatening their male egos.

In the *no threat condition Ss* were led to believe that *E* was interested only in the relationship between MAP and RT.

Apparatus.—MAP's were recorded by means of a Grass 8-channel electroencephalograph using bipolar surface electrodes. Copper electrodes, approximately 12 mm. in diameter, were placed on the ventral surface of the right and left forearms in accordance with recommendations made by Davis (1952). The electrodes were placed by having *S* flex his index finger while his arm was extended with the palm face down. This procedure facilitated the location of the forearm muscle group which is active in finger flexion. Each electrode was held in place by an elastic band approximately 15 mm. in width. Electrode jelly was used to reduce contact resistance with the skin.

The RT apparatus was constructed specifically for this experiment and is described more fully elsewhere (Rossi, 1958). In brief, the apparatus consisted of two white stimulus lights situated approximately 18 in. from each other in a horizontal direction. A red warning light was situated midway between the stimulus lights. Two toggle switches were placed in grounded metal boxes 18 in. apart and attached to a board which was placed across the arms of *S*'s chair during the experiment. Each of these switches controlled one of the stimulus lights. The positions of the stimulus lights could be manually reversed by *E* so that at times *S*'s right switch controlled the light to his right (simple task), and at other times *S*'s right switch controlled the light to his left (complex task). The left switch controlled the other stimulus light.

Procedure.—After the electrodes were attached to *S* he was asked to relax as much as possible while *E* calibrated the EMG. If *S* were to receive the simple-complex order of tasks, he was given four errorless practice trials on the simple task after which the experimental trials on the simple task were administered in the sequence L–R–L–L–R–L–R–R–L–R until 10 errorless trials had been recorded. The errorless trials on each task were not necessarily consecutive. The additional electrode was then attached, and the threat instructions were given to *S*. Four additional errorless practice trials on the simple task were given after which *S* received the instructions for the complex task. The experimental trials on the complex task were administered in the sequence R–L–R–R–L–

TABLE 1

ANALYSIS OF VARIANCE OF THE MAP SCORES FOR HI- AND LO-MAS GROUPS OBTAINED DURING TWO TASKS GIVEN UNDER TWO CONDITIONS

Source of Variation	*df*	*MS*	*F*
MAS	1	11,476.13	5.672*
Order (O)	1	165.02	0.082
MAS × O	1	120.12	0.059
Ss	32	2,023.25(e)	18.100**
Task (T)	1	171.13	1.531
Condition (C)	1	0.35	0.003
MAS × C	1	115.01	1.029
MAS × T	1	159.01	1.422
Pooled residual	32	111.78(e)	
Total	71	1,134.19	

Note.—The (e) indicates the error term used to obtain the *F* values reported above it.
* $P < .02$.
** $P < .001$.

R–L–L–R–L until 10 errorless trials had been recorded. There were 15 sec. between stimuli and 5 sec. between onset of warning light and onset of stimulus light. An identical procedure was followed with *Ss* who received the complex-simple order of tasks with the appropriate modification in the order of tasks.

Measurements.—The onset and offset of the stimulus lights produced pips on the MAP records, and the RT scores were obtained by measuring the distance in millimeters between the pips. Since the paper speed was 60 mm./sec., these scores were converted to milliseconds. Each *S*'s RT scores were averaged over the 10 errorless trials for each task.

The MAP records were scored by measuring the maximum deflection of the pens (in millimeters) during the 0.1-sec. interval immediately preceding *S*'s response. By reference to the calibration settings for each *S*, the millimeter score was converted to a microvolt score. These measurements were obtained only for the arm which was active in making the response. A logarithmic transformation of the microvolt scores yielded a more normal distribution than the raw microvolt scores. Accordingly, the statistical analyses are based on 100 log mv. scores. Each *S*'s MAP scores were averaged over the 10 errorless trials for each task.

RESULTS

MAS and MAP.—An analysis of variance was carried out on the MAP scores for the two MAS groups. The results of this analysis are presented in Table 1. The test of the MAS effect indicates that the two MAS groups had significantly different MAP scores. The mean MAP scores obtained during the simple task (S), simple task with threat involved (ST), complex task (C), and complex task with threat involved (CT) are presented in Table 2. An inspection of these means shows the HI–MAS group had consistently higher MAP scores throughout all tasks.

Also, it can be seen from Table 1 that the interaction effects between MAS and Condition did not reach an acceptable level of significance; there were no significant differences in MAP scores between the threat and no threat conditions for either of the MAS groups.

MAS and RT.—An analysis of

TABLE 2

MEAN MAP SCORES FOR HI- AND LO-MAS GROUPS OBTAINED DURING TWO TASKS GIVEN UNDER TWO CONDITIONS

MAS	Simple Task		Complex Task	
	No Threat	Threat	No Threat	Threat
HI–	247.22	255.22	246.78	243.56
LO–	224.11	221.89	224.44	221.33

TABLE 3

MEAN RT SCORES FOR HI- AND LO-MAS GROUPS ON TWO TASKS GIVEN UNDER TWO CONDITIONS

MAS	Simple Task		Complex Task	
	No Threat	Threat	No Threat	Threat
HI–	346.53	331.67	228.55	252.44
LO–	319.67	351.91	270.89	258.29

variance, similar to the analysis in Table 1, was carried out on the RT scores for the two MAS groups. A test of the Task effect in this analysis yielded an *F* ratio of 186.14 which is significant beyond the .001 level; the two tasks were significantly different in level of difficulty. The interaction of MAS with Task yielded an *F* ratio of 4.86, significant beyond the .05 level. An inspection of the means in Table 3 indicates that the HI–MAS group was superior on the simple-no threat task and inferior on the remaining tasks when compared to the LO–MAS group.

MAS, MAP, and RT.—Two additional analyses of variance, also similar to the analysis in Table 1, were carried out on the RT scores. In these two analyses the MAS effect was replaced by a MAP effect. One analysis employed only HI–MAS *Ss*, and the other analysis employed only LO–MAS *Ss*. Within each MAS group the HI– and LO–MAP groups were selected in the following manner. The 9 *Ss* who had received the simple-complex order of tasks were ranked according to averaged MAP score. The 4 *Ss* who had the lowest MAP scores were assigned to the LO–MAP group, and the 4 *Ss* who had the highest MAP scores were assigned to the HI–MAP group. The data from the remaining *S* were discarded because no criterion could be formulated for assigning him to either the HI– or LO–MAP group. A similar procedure was followed for the 9 *Ss* in each MAS group who had received the complex-simple order of tasks.

In the analysis of the RT scores for the LO–MAS MAP groups, an *F* ratio of 5.42 was obtained for the MAP effect, which is significant beyond the .05 level. With the exception of the Task effect, which was significant beyond the .001 level, none of the other *F* ratios reached an acceptable level of significance. The RT score means for the LO–MAS *Ss* with high MAP scores were higher on each task than the means for the LO–MAS *Ss* with low MAP scores.

In the analysis of the RT scores for the HI–MAS MAP groups, only the *F* ratio for the Task effect reached significance. It appears that for HI–MAS *Ss* high or low MAP scores are not consistently related to performance.

Discussion

In the present study it was found that the MAP scores for the groups did not vary with the nature of the task or the conditions under which the task was administered. Although the threatening instructions had no apparent effect on the MAP scores, the results on MAS and RT indicate that the instructions did have some effect on the performance scores. It will be recalled that the HI-MAS group was superior to the LO–MAS group on the simple-no threat task and inferior on the simple-threat task. Therefore, it appears that the threatening instructions had some effect on *Ss*' performances, but this effect was not due to a change in drive level. This finding suggests that the change in performance was associated with a change in the stimulus situation. Such an interpretation will be supported by the discussion which follows concerning what is measured with the MAS.

The HI–MAS group used in the present study had consistently higher MAP scores than the LO–MAS group. Accepting MAP scores as a measure of drive, this finding indicates that the HI–MAS group had the higher level of drive. This, in turn, indicates that the MAS has validity as a measure of drive. However, the other results of this study suggest that the MAS measures other factors in addition to drive level. These additional factors can best be described

by reference to the studies which have reported a high relationship between the MAS and the Psychasthenia scale from the MMPI (Ericksen, 1954; Ericksen & Davids, 1955). The Psychasthenia scale was constructed with items purported to measure ". . . mild depression, excessive worry, lack of confidence, or inability to concentrate" (Hathaway & McKinley, 1951, p. 20). The high correlations which have been reported between these two scales suggest that the MAS may also be a measure of these characteristics.

The results on MAS and RT also suggest that the MAS may be a measure of these characteristics. It will be recalled that the HI-MAS group was superior only on the simple-no threat task. The performance of this group became inferior on the remaining tasks. If the MAS is a measure of high drive *plus* "lack of confidence" and "inability to concentrate," then it may be expected that on a straightforward task with no distractions, high MAS scores would be associated with superior performance. It would also be expected that when the task became complex or when a diverting factor was introduced, performance of HI–MAS *Ss* would become inferior.

This interpretation makes the assumption that the threatening instructions did not raise the level of drive but rather introduced an extraneous factor into the task thereby making it complex. That is, it is assumed that the instructions to "react faster than the girls or get shocked" activated the tendency for HI–MAS *Ss* to ruminate and worry about their competency to perform well. The rumination then interfered with performance.

The results on MAS, MAP, and RT tend to confirm the hypothesis that the MAS is a measure of learned reaction patterns. Higher MAP was associated with superior performance for LO–MAS *Ss*. This finding is in agreement with the results of previous studies which have investigated the relationship between MAP and performance. However, no consistent relationship was noted between MAP and performance for the HI–MAS *Ss* in this study. This lack of relationship suggests that HI–MAS *Ss* have reaction patterns which interfere with the utilization of drive in a constructive manner on many tasks.

One of the purposes of the present study was to evaluate the hypothesis that the MAS is a measure of susceptibility to drive. The results on MAS and MAP disclosed no significant differences between the two MAS groups in reaction to the threat, which implies that the MAS does not measure susceptibility to drive. Therefore, to the extent that the scale does measure drive, it measures chronic drive. However, the differences in reaction patterns between HI– and LO–MAS groups may overshadow the differences in drive level between these groups.

Summary

This study was designed to evaluate the following alternative hypotheses: (*a*) the Manifest Anxiety Scale (MAS) is a measure of chronic drive; (*b*) the MAS is a measure of susceptibility to drive; (*c*) the MAS is a measure of learned reaction patterns to drive. The evaluation made use of electromyographic recordings (MAP) as an independent measure of drive states.

Two reaction time tasks (simple and complex) were administered to 36 *Ss* (18 LO-MAS and 18 HI-MAS) under two conditions (threat and no threat). Electromyographic recordings were taken from the arm active in making a response.

The results indicated that the MAS is a better measure of chronic drive than of susceptibility to drive. The patterning of results, however, indicated that the MAS also measures learned reaction patterns to drive which may overshadow its usefulness as a measure of chronic drive in any but simple tasks.

References

Davis, J. F. *Manual of surface electromyography.* Montreal: Lab. for Psych. Stud., Allan Mem. Inst. of Psychiat., McGill Univer., 1952.

Duffy, E. The psychological significance of the concept of "arousal" or "activation." *Psychol. Rev.*, 1957, **64**, 265–275.

ERICKSEN, C. W. Some personality correlates of stimulus generalization under stress. *J. abnorm. soc. Psychol.*, 1954, **49**, 561–565.

ERIKSEN, C. W., & DAVIDS, A. The meaning and clinical validity of the Taylor Manifest Anxiety Scale and the Hysteria-Psychasthenia Scales from the MMPI. *J. abnorm. soc. Psychol.*, 1955, **50**, 135–137.

FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln, Nebr.: Nebraska Univer. Press, 1954. Pp. 1–46.

HATHAWAY, S. R., & MCKINLEY, J. C. *Minnesota Multiphasic Personality Inventory manual*. New York: Psychol. Corp., 1951.

JESSOR, R., & HAMMOND, K. R. Construct validity and the Taylor Anxiety Scale. *Psychol. Bull.*, 1957, **54**, 161–171.

MALMO, R. B. Anxiety and behavioral arousal. *Psychol. Rev.*, 1957, **64**, 276–287.

MARX, M. Some relations between frustration and drive. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln, Nebr.: Nebraska Univer. Press, 1956. Pp. 92–130.

ROSSI, A. M. An evaluation of what is measured by the Taylor Manifest Anxiety Scale by the use of electromyography. Unpublished doctoral dissertation, Univer. of Utah, 1958.

SHORE, M. F. Perceptual efficiency as related to induced muscular effort and manifest anxiety. *J. exp. Psychol.*, 1958, **55**, 179–183.

SPENCE, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

TAYLOR, J. A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–320.

(Received August 14, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

RESPONSE LEARNING IN PAIRED-ASSOCIATE LISTS AS A FUNCTION OF INTRALIST SIMILARITY [1]

BENTON J. UNDERWOOD, WILLARD N. RUNQUIST, [2] AND RUDOLPH W. SCHULZ

Northwestern University

Paired-associate learning of verbal lists can be divided logically into two phases. These two phases will be called the response-recall or response-learning phase, and the associative phase. The response-learning phase is conceived of as the learning required to make the responses readily recallable. This first phase must necessarily precede the second phase, the associative phase, since the response must be available before it can be associatively connected (second phase) to a specific stimulus in the paired-associate list. It is possible, of course, that certain components of the response (e.g., the first letter) may develop associative strength to the stimulus before the entire response is available.

This division of rote learning into two phases is not new (e.g., Hovland & Kurtz, 1952), but so far as is known it has not been used systematically as an analytical device. Certain considerations suggest that variables which are known to affect the over-all rate of learning operate differentially during the two phases. For example, the fact that meaningfulness of responses in paired-associate learning influences the rate of learning to a greater extent than does stimulus meaningfulness (Sheffield, 1946) could be interpreted to indicate that as response meaningfulness increases both the response-recall phase and the associative phase are enhanced, whereas, when stimulus meaningfulness increases through a corresponding range only the associative phase is facilitated.

The present studies are concerned with intralist response similarity and its role in the two phases of learning as outlined above. The general hypothesis to be tested is that the higher the intralist response similarity the greater the facilitation of the response-learning phase. Since it is known (Feldman & Underwood, 1957) that with the type of lists to be used high intralist response similarity impedes over-all paired-associate learning (as compared with low intralist response similarity), it must follow from the hypothesis that the negative or interference component of high response similarity occurs primarily in the associative phase.

That the response-recall phase would be enhanced by high intralist response, similarity could be expected on at least two somewhat independent bases. First, it is known that high inter-item similarity (at least meaningful similarity) and associative connection are almost perfectly correlated (Haagen, 1949). Thus, if *S* recalls one response, the recall of other similar responses should be facilitated. The clustering effect of items of high similarity shown by Bousfield and Cohen (1955) suggests the operation of this principle. A second way of viewing the situation would suggest simply that if all responses in a paired-associate list are similar, *S* merely has to remember a single con-

[1] This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and the Office of Naval Research.

[2] Now at Hobart College.

cept and then give instances of that concept in the response-recall phase. If all responses are dissimilar, however, as many different "concepts" as there are responses must be learned. Either way of viewing the situation suggests that the response-recall phase will take place more rapidly with items of high similarity than with items of low similarity.

In the major conditions of the present studies two 10-item paired-associate lists are used. In one list the 10 responses are dissimilar adjectives, and in the other they are adjectives all of which have a more or less common meaning. In one set of conditions, *S*s are taught the 10 responses *before* they become responses in a standard paired-associate list. Teaching *S* the responses may be thought of as isolating the response-learning phase. The hypothesis stated above leads to two predictions about the results of these procedures. First, when *S*s are taught the responses before they become a part of a paired-associate list, it must be predicted that acquisition of the 10 similar responses will take place more rapidly than the acquisition of the 10 dissimilar responses. Secondly, having taught *S* the responses before they become a part of a paired-associate list, it must be predicted that when they do become responses in this paired-associate list the initial learning process will occur more rapidly than will learning in control conditions where the initial learning of the responses has not taken place. This prediction must be limited to the early trials corresponding to the period during which response learning is believed to occur. While it might appear that if response learning is facilitated the associative phase would "get started sooner" and that over-all learning would thereby be facilitated, there is nothing in the hypothesis which asserts this. Teaching *S* the responses before these responses become a part of a paired-associate list is assumed merely to remove the bulk of the response-learning phase in paired-associate learning.

In another set of conditions standard paired-associate learning is used but *S*s are stopped at various points in learning and are asked to give all the response items they can remember. The prediction is that more responses will be given from the high-similarity list than from the low-, (at least early in learning) even though the over-all rate of learning is slower for the high-similarity list.

METHOD

Materials.—For the basic learning task two lists of 10 pairs were used in which the stimulus items were nonsense syllables and the response items were two-syllable adjectives. The stimuli were taken from a previous study (Underwood, 1953a) and were of from 46–53% association value (Glaze, 1928) and of low intralist similarity. The responses were also taken from a previous study (Underwood & Goad, 1951). In one list all responses were more or less synonymous in meaning (*cheerful, sunny, carefree, pleasant, gleeful, laughing, happy, genial, jolly, smiling*). In the other list the responses were unrelated in meaning (*spicy, rounded, hairy, equal, modern, tiresome, fiery, faithful, plastic, guilty*). On the average both sets of responses have about the same Thorndike-Lorge (1944) frequency. In pairing a syllable and an adjective an attempt was made to avoid any apparent connections between the two.

General design.—Three basic conditions each involved paired-associate learning of the high-similarity list and the low-similarity list, with separate groups learning each list in each condition. In one condition of response learning (Cond. RL), *S* learned the responses alone before learning the paired-associate list. In this RL procedure the responses were presented in varying order from trial to trial for five trials with a recall test being given after each trial. One group learned the high-similarity responses and another the low-similarity responses before proceeding to their respective paired-asso-

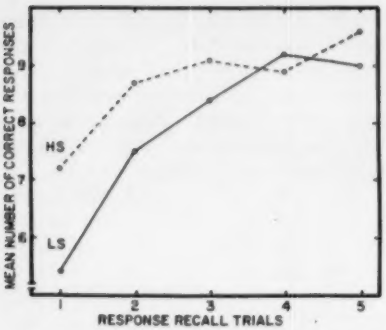

FIG. 1. Mean correct responses on each response-recall trial as a function of high similarity (HS) and low similarity (LS).

ciate lists. A control condition (Cond. C) involved no pretraining. A comparison of learning the paired-associate lists for RL and C conditions should indicate the effectiveness of learning the responses prior to paired-associate learning. The third condition (Cond. PA-RL) was designed to determine the extent of response learning (RL) during paired-associate learning (PA). Various groups were stopped at different points during paired-associate learning and given a single free recall of the responses. This response learning was tapped at six points, namely, after 1, 2, 3, 5, 8, and 13 exposures of the paired-associate lists. Since each *S* was stopped only *once* during learning, the PA-RL conditions involved 12 groups, six for each of the paired-associate lists.

Subjects and procedure.—A total of 320 *S*s served in the experiment, 20 in each of the 16 groups. Condition RL involved 2 groups; Cond. C, 2 groups; and Cond. PA-RL, 12 groups. The *S*s were assigned to one of the 16 groups randomly with the restriction that one *S* was assigned to each group before there were any replications. All *S*s were naive to verbal learning experiments.

In the standard paired-associate learning 15 anticipation trials were given under all conditions. Lists were presented on a memory drum at a 2/2-sec. rate, with a 4-sec. intertrial interval. Five different orders of the pairs were used to minimize serial learning with an equal number of *S*s being started on each order.

For the response-learning phase of the two RL groups, the list of responses was presented in five alternate learning and recall trials. The *S* was clearly instructed that he was not to learn the adjectives in order since they would be in a different order on each trial. The recall period after each learning trial was 1.5 min. The *S* simply wrote down on a sheet of paper as many responses as he could remember. Since there were five trials and five orders, an equal number of *S*s was started on each order to avoid any possible serial position bias. None of these orders was the same as subsequently used in paired-associate learning. For the paired-associate learning for these RL groups the *S* was provided with a card on which all the responses were listed. The intent of this was that if any of the responses were forgotten of those just learned, or if learning had been incomplete on five trials, the responses would still be available. However, few *S*s appeared to use the information on these cards; the majority reported that they simply did not have time to make use of the information during paired-associate learning.

In the PA-RL conditions, *S* was stopped after the appropriate number of trials on the paired-associate list, read the instructions for response recall, and then given 1.5 min. to write down all the responses he could remember. The *S* did not know beforehand that such a recall would be asked for. The instructions made it clear to *S* that he was to write down all the adjectives he could remember without regard to order. After the recall period he was informed that he would not be interrupted again and the standard learning of the paired-associate list continued until 15 anticipation trials were completed.

RESULTS

Response learning.—The mean number of correct responses on each of the five response recall trials for the RL groups is shown in Fig. 1. Initially the similar responses are recalled better than the dissimilar responses, although by Trial 4 performance is approximately equal on both lists and asymptotic for a majority of *S*s in both groups. The t on the total correct responses over all five trials for the two groups was 3.09 ($P < .001$). Thus, relatively pure response learning occurs more rapidly for responses of high similarity than for those of low similarity.

The results from the response recall trials for the PA-RL groups are plotted in Fig. 2. A comparison of Fig. 1 and Fig. 2 shows that response learning is much higher when occurring independently (Fig. 1) than when measured as a component of standard paired-associate learning (Fig. 2). Nevertheless, the same general relationship between the recall of high- vs. low-similarity responses still obtains. The curves are again initially separated, with recall best for high-similarity items, although the performance becomes equivalent after only a few trials. An analysis of variance was performed using data from the groups with 1, 2, 3, and 5 previous learning trials (since the prediction concerns only initial learning), making a 2 × 4 factorial design with trials and similarity being the two variables. The *F*'s for trials (21.18, 3 and 152 *df*) and similarity (10.51, 1 and 152 *df*) are both significant beyond the .01 level, while the interaction is not significant ($F = 1.32$).

It can be shown that response learning measured during paired-associate learning (Fig. 2) is considerably higher than over-all paired-associate

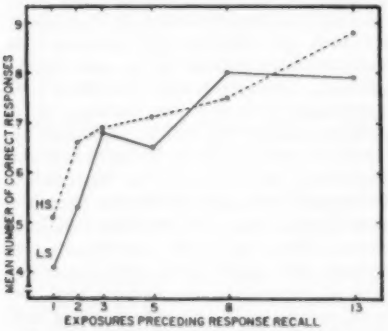

FIG. 2. Mean correct responses on response recall during paired-associate learning as a function of similarity (HS vs. LS) and number of exposures preceding response recall.

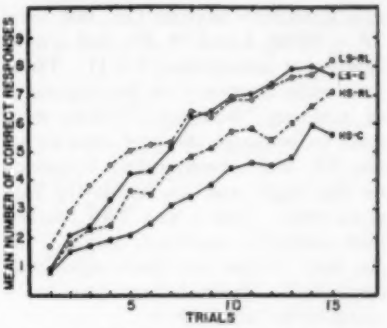

FIG. 3. Paired-associate learning in Exp. I as a function of preresponse learning (RL vs. C) and intralist response similarity (HS vs. LS).

learning as such. While this would be expected from the general conception being evaluated (response learning precedes the associative phase), the data of the present experiment do not allow a satisfactory conclusion concerning this matter. This is because in the measurement of response recall during paired-associate learning essentially unlimited recall time was given, while in paired-associate learning only 2-sec. per item were allowed.

Paired-associate learning.—The acquisition curves during paired-associate learning for the RL and C groups for both high- and low-similarity lists are presented in Fig. 3. In both kinds of lists there is evidence of facilitation due to the prior learning of the responses. The facilitation in the high-similarity list is apparent throughout the 15 trials, while facilitation in the low-similarity list is restricted to early trials. Again, since the prediction only involves the early trials, statistical evaluation of these results was made on the total correct responses over the first six trials. Analysis of variance showed that response learning was significant beyond the .02 level ($F = 6.35$, 1 and 76

df), similarity beyond the .001 level ($F = 16.00$, 1 and 76 *df*), and a nonsignificant interaction ($F < 1$). There was some evidence of heterogeneity of variance; however, *t* tests were used to compare the performance on the RL and corresponding C groups on the high- and low-similarity lists separately. The *t* was 1.90 for the high-similarity condition and 1.77 for the low. These are both significant beyond the .05 level if one assumes a single-tailed alternative.

Other evidence on response learning.—From the above comparisons it may be inferred that response learning of similar items occurs more rapidly than response learning of dissimilar items. Thus, early in learning a paired-associate list, *S* should have more responses available if the responses have high similarity than if they have low similarity. If this is true, evidence for it might be found in actual paired-associate learning. The point at which such evidence would be least contaminated by the associative phase would be the first anticipation trial. The *S* would be expected to give more *different* responses on this trial if the responses have high similarity than if they have low similarity. Such an analysis ignores the question of whether the response was appropriately paired with its stimulus; rather, a simple count is made of the number of different responses occurring. Five such comparisons are available from the PA-RL conditions; that is, there are five conditions in which a low-similarity list was taken through at least one anticipation trial by standard procedures and five comparable conditions for high-similarity lists. There is an additional comparison which can be made using the C groups, and one further comparison from a second experiment to be presented shortly. The comparisons were between the mean number of different responses given on the first anticipation trial of the high-similarity and low-similarity lists. In six of the seven comparisons, the mean was larger for the high-similarity lists. The over-all mean was 2.31 for the high-similarity lists and 1.76 for the low-similarity lists.

In contrast to the above expectations, if *S* has been taught the responses before paired-associate learning (as was true in the RL conditions) the high-similarity list should not show a superiority in number of different responses. Two tests of this can be made, one from the present RL groups and another from two comparable groups appearing in Exp. II. In both of these comparisons the mean number of different responses was higher for the low-similarity lists than for the high-similarity lists.

Discussion.—The prediction that response recall will be initially greater when the responses are similar in meaning than when they are dissimilar is clearly supported in all of the above data. This is true whether the response recall is taken during paired-associate learning, whether a response-learning procedure is used, or whether the responses in paired-associate learning as such is examined. The prediction that the initial performance during the learning of a paired-associate list will be facilitated by the prior learning of the responses is not supported as clearly. The differences, while small, are in the predicted direction and closely approach statistical significance. However, it is possible to attribute this small facilitation to some kind of warm-up or learning-to-learn (or both) produced by response learning and which is independent of the particular responses which were learned. Therefore, Exp. II was run which duplicated the RL and C conditions with the exception that the C groups in this sec-

ond study were given irrelevant response learning prior to paired-associate learning. Thus, if relevant response learning facilitates paired-associate learning over and above that produced by irrelevant response learning, the differences found in Exp. I cannot reasonably be attributed to warm-up and practice effects.

Experiment II

Method.—The RL conditions of the previous experiment were replicated exactly, with both the high- and low-similarity lists being used. The C groups in Exp. II, however, were given five response learning and recall trials on irrelevant lists before being given the paired-associate lists. The lists were irrelevant in the sense that none of the items appeared in the paired-associate list. Such irrelevant learning should produce the same warm-up and practice which the RL groups might get in learning a relevant list. The irrelevant responses learned in the response-learning phase for the group subsequently learning the paired-associate list with high similarity among responses were: *shifty, crafty, expert, cunning, cagey, wily, foxy, clever, skillful,* and *tricky.* For the group learning the paired-associate list with low similarity among responses the items for the irrelevant response learning were: *hybrid, crumbling, vulgar, fiscal, worldly, inform, warlike, neuter, sterile,* and *flashy.*

A total of 60 *Ss* served in the experiment, 15 in each group. All *Ss* were naive to verbal learning experiments. All procedures were exactly the same as those in Exp. I, except as noted above, and except for the fact that during paired-associate learning the RL groups were not provided the card on which the correct responses were written.

Results.—The response-recall data showed essentially the same results as did Exp. I. The difference between high and low similarity was somewhat less than in Exp. I, but still significant statistically on the first recall trial.

As may be seen in Fig. 4, the results of the learning of the paired-associate list duplicate almost exactly those of the first study. For both high- and low-similarity lists the RL group exhibits better performance than the C group. The only clear difference

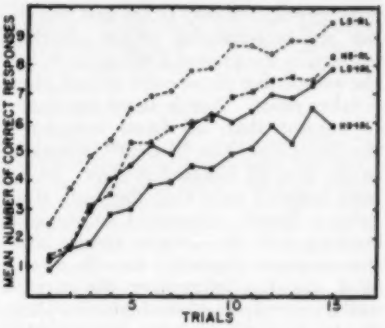

Fig. 4. Paired-associate learning in Exp. II as a function of preresponse learning (RL vs. IRL) and intralist response similarity (HS vs. LS).

between the results for the two studies is that in the present results the RL group is facilitated throughout all 15 trials on the low-similarity list; this was not true in the first study. Direct comparisons between the results of the two studies are not justified since the *Ss* in Exp. II, as a group, were faster learners than those in Exp. I.

Analysis of variance of the total correct responses on the first six trials gave significant *F*'s for both similarity (7.53) and relevance of response learning (9.81). With 1 and 56 *df*, the *F* at the .01 level is 7.12. Separate *t* tests between the RL and C groups gave values of 2.32 for high similarity and 2.23 for low similarity, both of which are significant at the .05 level with 28 *df*. The conclusion seems justified that prior learning of the responses facilitates paired-associate learning over and above any nonspecific effects, such as warm-up or learning-to-learn.

Discussion

The results are consonant with a two-phase conception of verbal learning. It should be clear that the present results are

directly relevant only to the first phase—the response-learning phase. Nothing is given in the present findings as to how the associative phase—the second phase—takes place. Nor is there any intent to suggest that, all items considered, the two phases can be clearly separated in the over-all learning process. It has been asserted only that the early trials have a heavy component of response learning and the present results show that response similarity may be viewed as a variable influencing the rate at which this response learning takes place. Furthermore, the results indicate that if the necessity for response learning in paired-associate learning is removed by teaching *S* the responses before they become a part of the paired-associate list, paired-associate learning is facilitated. Concerning this latter finding it should be noted that Hovland and Kurtz (1952) have shown that familiarization training will facilitate serial learning. However, serial learning tends to confuse stimulus and response functions of the items. While stimulus learning before paired-associate learning (given in the same manner that response learning was given here) might influence subsequent paired-associate learning, it would be through other mechanisms (e.g., stimulus differentiation) than those being dealt with here for response learning.

It might be suggested that the response-learning procedures used produced a differentiation among the responses (reduced intralist similarity), hence facilitated learning. A number of facts argue against this interpretation. (*a*) If differentiation among responses is involved, response learning of the lists of low similarity should have little or no effect in paired-associate learning. (*b*) If differentiation among responses is the critical factor, it should have produced a much greater facilitation in the learning of the paired-associate lists for responses with high similarity than for responses with low similarity. The results showed that there was no appreciable difference between the effect for responses of low similarity and responses of high similarity. In fact, the mean differences for the initial part of learning favored the lists with low response similarity. (*c*) Finally, there is no reason to believe that the *method* used for response learning would produce a differentiation among responses in the sense that associations were differentially reinforced. All facts considered, therefore, it does not seem likely that response differentiation is of much moment in the present studies.

The present materials probably do not allow an optimum effect to be produced by response learning. That is, since adjectives were used, response learning should take place much more rapidly than if, say, nonsense syllables of low-association value had been employed. (It will be recalled that pure response learning in the RL groups took place very rapidly.) It would be expected that if response learning were given before paired-associate learning with material of low meaningfulness, subsequent paired-associate learning would be facilitated more than was true in the present results.

What is the stimulus for the recall of each item in the response-learning phase? There is no satisfactory answer to this question. But, it should be pointed out that the question is relevant not only to the present procedures but also to any learning situation in which *S* is shown a group of items and is then asked to recall them in any order. The stimulus for such recall may be the gross stimulus complex of the laboratory room, the memory drum, the instructions, etc. No resolution of this systematic problem seems possible. At the same time it must not be inferred that the initial recall of a given response is never instigated by the particular stimulus with which it is paired in the list. The fact that it is possible to set up pairs in which already established associations obtain between the stimulus and response, and that learning is very rapid in such situations, suggests that response recall can be tied to the specific stimulus and that this is a very effective arrangement. It is therefore quite possible that in the present experiment the initial recall of

some items by some *Ss* is entirely tied to a specific stimulus.

The present findings on the role of intralist response similarity have some implications for the role of both stimulus and response similarity in over-all paired-associate learning. Although available data are not completely conclusive on this matter (Underwood, 1953a; 1953b), the best evidence is that a given range of intralist stimulus similarity produces a greater variation in learning than does a corresponding range of response similarity. Such a finding is understandable in light of the present results and the conception of two stages in learning. It could be assumed that increasing interference produced by increasing similarity would affect the associative phase of learning and would have an equal effect in this phase for stimulus and response similarity. But, increasing response similarity has an increasing positive effect for response learning in the first phase (as seen in the present results) which has no counterpart in stimulus similarity. Therefore, the over-all effect of stimulus similarity should be greater than the effect of response similarity.

Summary

Verbal learning may be conceptualized as a two-stage process. In the first phase *S* must learn the responses, in the second he must attach them to specific stimuli. The present experiments dealt directly with the first stage only. It was hypothesized that: (*a*) response learning is initially more rapid the higher the response similarity in a paired-associate list; and (*b*) teaching *S* the responses before he learns a paired-associate list would initially facilitate the learning of this list.

Two paired-associate lists were used. The stimuli for both lists were nonsense syllables, the responses, adjectives. In one list the adjectives were all similar in meaning, in the other, dissimilar. In one set of conditions (control), *Ss* merely learned the paired-associate lists by standard procedures for 15 trials. In a second set of conditions, *Ss* were taught the responses prior to learning the paired-associate list for 15 trials. In the third set of conditions standard paired-associate learning was used but different groups of *Ss* were stopped after 1, 2, 3, 5, 8, and 13 trials and were asked to write down all the responses they could remember. These three sets of conditions required 16 groups of *Ss*. There were 20 in each group, all naive to verbal learning experiments.

The results show that:

(a) Teaching *S* the responses prior to paired-associate learning facilitates the learning of lists with both high and low similarity among the responses. A second experiment showed that this effect cannot be ascribed to warm-up or learning-to-learn resulting from the procedure of teaching *S* the responses before paired-associate learning. In both experiments the positive effect was evident throughout the entire 15 trials for lists of high response similarity. For low similarity the effect was only in initial learning for one comparison but present throughout learning in the other.

(b) In response learning (prior to paired-associate learning) items with high similarity are learned initially more rapidly than are items of low similarity, although for both lists learning was very rapid. When *Ss* are tested for free recall at various points in learning a paired-associate list, more responses are given from a high-similarity list than from a low-similarity list even though over-all level of paired-associate learning is higher for the low-similarity list. The difference in this response recall was clearly evident for the first few trials.

The results confirm the expectation that high intralist response similarity would facilitate response learning. Thus, the results are consistent with the two-stage conception of learning. Furthermore, the present results aid in understanding certain previous findings on the roles of stimulus similarity and response similarity in verbal learning.

References

Bousfield, W. A., & Cohen, B. H. The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. *J. gen. Psychol.*, 1955, **52**, 83–95.

Feldman, S. M., & Underwood, B. J. Stimulus recall following paired-associate learning. *J. exp. Psychol.*, 1957, **53**, 11–15.

Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.

Haagen, C. H. Synonymity, vividness, familiarity, and association-value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **30**, 185–200.

HOVLAND, C. I., & KURTZ, K. H. Experimental studies in rote-learning theory: X. Pre-learning syllable familiarization and the length-difficulty relationship. *J. exp. Psychol.*, 1952, **44**, 31–39.

SHEFFIELD, F. D. The role of meaningfulness of stimulus and response in verbal learning. Doctoral dissertation, Yale Univer., 1946.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Columbia Univer. Press, 1944.

UNDERWOOD, B. J. Studies of distributed practice: VIII. Learning and retention of paired nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 133–142. (a)

UNDERWOOD, B. J. Studies of distributed practice: IX. Learning and retention of paired adjective lists as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 143–149. (b)

UNDERWOOD, B. J., & GOAD, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.

(Received August 14, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

THE EFFECT OF PRECISION, DELAY, AND SCHEDULE OF KNOWLEDGE OF RESULTS ON PERFORMANCE

F. J. McGUIGAN [1]

Hollins College

The purposes of this investigation were: (*a*) to determine the effect of variation of three variables, precision (specificity), delay, and schedule of knowledge of results (KR), on performance; (*b*) to determine whether interactions result from the simultaneous variation of these three variables; (*c*) to determine the effects of removing KR on later performance; and (*d*) to determine whether quadratic and cubic components of these conditions exist in addition to linear components. An attempt was made to answer these questions using the motor task of line drawing.

EXPERIMENT I

Method

Subjects.—The *S*s were 84 women students from introductory psychology classes at Hollins College.

Apparatus.—The apparatus for line drawing was a wooden drawing board with an L-shaped piece of wood nailed to its top. The long arm of the L was not nailed to the drawing board so that a piece of graph paper, which was previously marked to identify various lengths from the starting point, could slide under it. A manually operated electrical timer was used for the various delay operations.

Procedure.—The *S*s were instructed to draw 6-in. lines while blindfolded. After 70 trials in which KR was furnished according to the procedure discussed below, *S*s were allowed a 5-min. rest period. Following this they were again blindfolded and asked to draw 70 more lines, this time with no KR furnished them. The first 70 trials will be referred to as "learning trials" while the second 70 trials will be referred to as "extinction trials." The *S*'s hand always remained at the point where it stopped until returned to the starting point by *E* for the next trial. At no time was *S* allowed to see the apparatus or the lines drawn, or to touch the drawing board with her moving hand.

Design.—A $3 \times 3 \times 3$ factorial design was used. Each of the three variables (precision, delay, and schedule of KR) was varied in three ways. The *S*s were randomly assigned to cells, three to each cell. In addition, a control group composed of three *S*s was used. The control group was never given KR. The three values of the precision variable were 1/8 in., 5/8 in., and 10/8 in.[2] That is, *S*s were told within these units how much their lines deviated from 6 in. Thus, for the 1/8 in. condition, *S* was told that the line was "correct" if it was within 1/8 in. of 6 in.; if it was within 2/8 in. (but more than 1/8 in.) of 6 in. she was told "1/8 in. too long (or too short)," etc. For the 5/8-in. and 10/8-in. conditions, *S*s were told "correct" if their lines did not deviate more than these values from 6 in., and so forth.

The three conditions of length of delay of KR were: immediately (0 sec.), 15 sec., and 30 sec. after the response terminated. Thus, the procedure for an *S* under the 15-sec. delay condition would be: *S* draws a line, waits 15 sec., is given KR, her hand is then returned to the starting point, and a new line is drawn.

The three conditions for schedule of RK were: 10%, 55%, and 100%. For example, for the 10% condition, *S* was furnished KR on 10% of the 70 trials. The trials on which KR was furnished were randomly determined in the case of the 10% and the 55% schedules. If an *S* was on less than a 100% schedule and also in either the 15-sec. or the 30-sec. delay condition, she was required to delay the appropriate length of time after completing her response, even though no KR was given on that trial.

[1] The author wishes to express his appreciation for assisting with the data collection to Mary Carr (Exp. I), to Linda Henley and Frances Crockett (Exp. II), and to R. Lowell Wine for assistance with the statistical analysis.

[2] It is assumed that the results for these intervals do not differ from those for equal interval values.

TABLE 1

MEAN DEVIATIONS PER *S* AS A FUNCTION OF CONDITION AND STAGE

Condition	Stage													
	Learning							Extinction						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Delay (Sec.)														
0	93.1	71.8	62.5	55.2	50.5	49.0	56.1	73.3	67.6	73.5	75.7	65.5	69.3	67.1
15	91.9	54.0	54.1	44.1	40.3	42.0	44.8	71.3	66.7	70.9	67.2	68.3	73.9	72.1
30	98.3	60.4	55.0	46.2	48.4	42.1	46.0	54.6	46.4	44.3	51.8	53.7	47.8	58.7
Schedule														
100%	63.3	41.1	41.0	41.3	38.9	35.9	39.3	57.1	58.2	59.3	60.0	57.7	62.1	62.2
55%	86.9	51.1	43.9	39.2	44.1	41.4	42.6	62.6	54.2	56.0	55.3	58.4	59.7	61.7
10%	133.0	94.0	86.7	65.1	56.1	55.7	65.0	79.6	68.2	73.4	79.3	71.4	71.7	74.0
Precision														
1/8 in.	94.6	55.7	42.1	33.5	35.4	29.0	34.8	72.7	55.2	59.4	65.0	49.6	50.8	52.4
5/8 in.	81.6	64.3	67.7	49.9	42.0	45.9	45.6	57.4	63.2	64.9	65.7	70.0	70.6	72.1
10/8 in.	107.0	66.3	61.9	62.2	61.7	58.1	66.6	69.0	62.3	64.4	64.0	67.8	69.5	73.5

Results

The deviation from a line 6 in. in length was measured in units of 1/8 in. for each *S* for each trial. These deviation scores were summed in groups of 10 trials for each *S*, which resulted in 7 stages for the learning trials and 7 stages for the extinction trials. The mean deviation per *S* for each stage is presented for the several conditions in Table 1.

The data for learning and extinction trials were analyzed separately using the method of orthogonal polynomial analysis of trends for repeated measurements of the same *S*s, as suggested by Grant (1956). The results of this analysis for the learning trials are presented in Table 2 and will be discussed first. We shall restrict our discussion to *F*'s that are significant at the 1% level, although the reader may observe the probabilities of the other *F*'s in Table 2.

The significant *F* for over-all trend indicates that the mean scores of the 81 *S*s included in the factorial design vary from stage to stage. That is, the over-all trend for the 81 *S*s deviates from the horizontal line throughout the seven stages. A glance at Table 1 suggests a general negative slope for the over-all trend, which indicates significant improvement in performance as number of stages increase. That the over-all trend has significant linear and quadratic components is indicated by the significant *F*'s for each of these components. This means that there is a significant quadratic trend in addition to the significant linear trend.

Evaluation of the sources of variance of the between-group means is made in the usual way. The lack of a significant *F* for delay indicates that variations of this variable did not lead to significant differences in performance. That the three conditions of schedule of KR led to differences in performance is indicated by the significant *F* for this variable. From Table 1 it may be seen that the 100% schedule resulted in the best performance, followed next by the 55% condition, while the 10% condition led to the poorest performance. None of the interactions were significant. It may thus be seen that the between-group means variation is

mainly accounted for by the KR schedule and the precision variables.

The analysis of between-group trends indicates that schedule is the only variable for which there is a significant linear interaction with stages. That is, the extent of the superiority of a certain value of the schedule variable depends on the particular stage that one observes.

The significant F for between-individual means indicates that there are significant differences among the performances of individuals throughout stages. As Grant (1956) points out, this indicates that the measures of performance are reliable.

The interpretation of the F's in Table 2 may be better understood by studying Table 1; this is particularly appropriate for understanding the precise nature of the interactions.

The same analysis was conducted for extinction trials. The only F significant at the 1% level, however, was for between-individual means ($F = 14.82$), again an indication that the measures of performance were reliable. The following F's, however, approached significance: an F of 3.84 for the linear precision component of the between-group trends source of variance ($P < .05$); an F of 1.65 for between-group means ($P < .10$); and an F of 2.58 for the delay component of between-group means ($P < .10$) with the 30-sec. delay condition being superior. The failure of the F for over-all trend to be significant indicates that there is a lack of departure from a horizontal line for all curves, a fact that can be seen by observing performance under all conditions during the extinction period (Table 1).

It will be recalled that three control *Ss* were run under a condition of no KR during either the learning or the extinction trials. The performance scores of these *Ss* were compared to those of the 81 *Ss* who received KR. More precisely, the deviations from a 6-in. line were summed throughout the 70 learning trials and also throughout the 70 extinction trials for each *S* in both groups. The resulting t for the learning trials was 2.66 ($P < .02$), and for extinction it was 3.26 ($P < .01$), the control *Ss* being inferior in both cases. These findings indicate that KR, in general, led to superior performance both during the 70 trials when KR was furnished, and for the 70 trials when no KR was furnished.

Discussion

The finding that the higher percentage schedules of KR led to performance that

TABLE 2

ANALYSIS OF VARIANCE FOR LEARNING TRIALS

Source of Variation	*df*	*MS*	*F*	*P*
A. Over-all trend (Trials)	(6)	24,742.11	39.32	<.01
Linear	1	96,434.97	69.03	<.01
Quadratic	1	43,333.81	50.03	<.01
Cubic	1	3,839.42	6.19	<.02
B. Bet. group means	(26)	13,138.26	1.99	<.05
Precision	2	24,397.66	3.70	<.05
Delay	2	4,430.76	—	—
Schedule	2	70,584.34	10.70	<.01
P×D	4	11,688.26	1.77	<.20
P×S	4	3,102.18	—	—
D×S	4	6,673.16	1.01	>.20
P×D×S	8	7,114.35	1.08	>.20
C. Bet. group trends	(156)	823.86	1.31	<.05
1. Linear	(26)	2,289.42	1.64	<.10
Precision	2	2,601.89	1.86	<.20
Delay	2	248.94	—	—
Schedule	2	13,032.01	9.33	<.01
P×D	4	2,293.89	1.64	<.20
P×S	4	705.29	—	—
D×S	4	140.22	—	—
P×D×S	8	1,900.20	1.36	>.20
2. Quadratic	(26)	898.92	1.04	>.20
Precision	2	2,738.56	3.16	<.10
Delay	2	161.93	—	—
Schedule	2	3,020.47	3.49	<.05
P×D	4	534.89	—	—
P×S	4	1,212.86	1.40	>.20
D×S	4	274.54	—	—
P×D×S	8	430.11	—	—
3. Cubic	(26)	891.57	1.44	<.20
Precision	2	2,134.62	3.44	<.05
Delay	2	766.08	1.23	>.20
Schedule	2	1,434.23	2.31	<.20
P×D	4	1,508.53	2.43	<.10
P×S	4	836.66	1.35	>.20
D×S	4	580.17	—	—
P×D×S	8	351.18	—	—
D. Bet. Ind. Means	54	6,593.99	10.48	<.01
E. Bet. Ind. Trends	(324)	629.22	—	—
1. Linear	54	1,397.00	—	—
2. Quadratic	54	866.16	—	—
3. Cubic	54	620.40	—	—
Total	566	2,082.18	—	—

was superior to the 10% schedule is in general agreement with the results obtained by Houston (1947). Using the Pedestal Sight Manipulation Test, he varied schedule in three ways: 100%, 50%, and 25%, and found that during learning, *Ss* on a 100% schedule performed significantly better than *Ss* on a 25% schedule, although no other differences were significant. The results reported by Goldstein and Rittenhouse, who used 100%, 50%, and 0% schedules with the Pedestal Sight Manipulation Test and the Flexible Gunnery Research Device, appear contrary to the present findings, for they conclude that this variation "made little difference" as far as performance is concerned (1954, p. 196).[3]

It is interesting to compare the results of varying the schedule of KR with those that result from varying the schedule of reinforcement. Thus, in the latter type of study, Jenkins and Stanley report that "Response strength is built up somewhat more rapidly under a schedule of 100% reinforcement than under a partial regimen . . ." and that "In almost every experiment, large and significant differences in extinction favoring the groups partially reinforced in conditioning over the 100% ones were found" (1950, p. 231). While the first generalization seems applicable in the case of KR, the second does not (cf. Bilodeau & Bilodeau, 1958).

The finding that variation of the delay of KR did not affect performance need not surprise us, for previous results with regard to this variable are contradictory. Thus while several studies report that the longer the delay, the poorer the performance (cf. Ammons, 1956; Bourne, 1957; Greenspoon & Foreman, 1956, McGuigan, Crockett, & Bolton, in press;) other findings are to the contrary (e.g., Bilodeau & Bilodeau, 1958; Johnson, 1956; Saltzman, Kanfer, & Greenspoon, 1955). Where the effect of the delay variable is studied using a repetitive task, three variables are involved: delay of KR after a response, time interval by which KR precedes the next response, and intertrial interval. Since these three variables are not independent, it is necessarily the case that at least two of them are confounded. Thus the confounded variables may well assume different values in different experiments on this variable, leading to contradictory results. The impossibility of unambiguously determining the effects of the length of time that KR follows a response in a repetitive task is discussed elsewhere (McGuigan, 1959).

Previous studies indicate that the withdrawal of KR leads to decreased performance (cf. Ammons, 1956). Houston (1947) found that when KR was removed, his various groups failed to differ significantly on extinction trials. Similarly, Goldstein and Rittenhouse (1954) found that whatever advantages had accrued during training with KR were essentially lost by the end of 20 extinction trials. While performance in the present study was poorer when KR was removed, the finding that the no-KR control group was significantly inferior to the pooled scores of the 27 experimental groups appears to be in direct contrast to the previous results. One is inclined to ask, however, whether this significance would be lost if extinction trials were continued beyond the 70 given in the present experiment.

Finally, one cannot help but be impressed by the sizable amount of information that can be obtained with the statistical analysis described by Grant (1956). It is to be expected that this type of analysis will be more widely used in our future experimentation.

Experiment II

Aside from the confounding of the length of delay of KR after a response, as discussed above, a possible reason for the negative results for this variable

[3] We are relating our results to those of previous experiments, even though the tasks differ. It is to be hoped that principles of KR may be generalized to various populations of tasks. In the event that interactions are discovered between conditions of KR and task characteristics, such generalizations, of course, will have to be restricted.

in Exp. I is that *S*'s hand was returned to the starting point for a new trial as soon as KR was furnished. In general, Greenspoon and Foreman (1956) had their *S*s wait varying periods of time after KR was provided before returning *S*'s hand to the starting point. It is possible that this difference in procedure partially explains the difference in results, for their experiment resulted in positive findings (Greenspoon & Foreman, 1956). That is, it is possible that the activity of moving the hand immediately after KR was furnished retroactively interfered with the effectiveness of the information. To test this hypothesis, a second experiment was run using 64 *S*s from the same population used for Exp. I. The same general procedure as before was followed except that *S*s were given only 50 learning trials and 50 extinction trials and that they were instructed to draw 3-in. lines. A $2 \times 2 \times 2$ factorial design was used in which the length of delay of KR after the response was varied in two ways: (*a*) KR was furnished immediately after each response; (*b*) 10 sec. after the response. The two conditions of the second variable were as follows: (*a*) the hand was returned to the starting point immediately after KR was furnished, following which *S*s waited 10 sec. before making the next response; (*b*) *S*'s hand remained where it stopped on any trial for 10 sec., following which it was returned to the starting point and a new trial was immediately commenced. The third variable was experimenter's.

The length of the lines for each *S* were measured in units of 1/16 in. and summed for learning and for extinction trials. Separate analyses of variance were then conducted for these two sets of trials. None of the *F*'s for the learning trials approached significance. An extremely interesting result, however, showed up for the extinction analysis. As is shown in Table 3, a significant difference was obtained between the conditions of removing the hand immediately after KR was furnished and of not removing it immediately. The mean length of the lines drawn during extinction for the group that waited 10 sec. before having their hand removed to the starting point was 3.00 in., while for the group that had their hand removed immediately it was 3.36 in. Quite clearly, then, we can conclude that removing the hand immediately after furnishing KR leads to poorer performance during extinction trials. Why the difference did not show up during learning trials is not clear, although it is possible that if these trials had been continued for another 50 trials the difference would have been manifested there. In any event these results suggest that motor activity can retroactively interfere with the effectiveness of verbal information, as shown by later performance when no KR is furnished.

TABLE 3

ANALYSIS OF VARIANCE FOR EXTINCTION TRIALS

Source of Variation	*df*	*MS*	*F*
Hand removal (H)	1	5111.21	8.73**
Experimenters (E)	1	2381.22	4.07*
KR delay (KR)	1	426.96	—
E × H	1	831.75	1.42
E × KR	1	574.56	—
H × KR	1	82.81	—
E × H × KR	1	305.56	—
Error	56	585.38	

* Significant beyond 5% level.
** Significant beyond 1% level.

SUMMARY

The purposes of this investigation were to determine the effect of variation of three variables, precision, delay, and schedule of KR on performance during 70 learning and 70 "extinction" trials. Eighty-one *S*s were assigned to cells of a $3 \times 3 \times 3$ factorial design and performed a line-drawing task. Three additional *S*s formed a no-KR control group. The positive findings for learning trials were: (*a*) There was a significant over-all trend of negative slope with significant linear and quadratic components, indicating improvement in performance as number of trials increased; (*b*) The greater the percentage of trials on which KR was given, the better the performance, although there was a significant linear interaction of schedule with trials. For both learning and extinction trials differences among individuals were

significant, indicating that the measure of performance was reliable. Comparison of the 27 KR groups with the control group showed that the former were superior for both learning and extinction.

One possible reason for the finding that the length of time by which KR follows a response did not affect performance is that *S*'s hand was returned to the starting point immediately after KR was given. It was hypothesized that this motor activity retroactively interfered with the effectiveness of the KR. A second experiment using 64 *Ss* was conducted to test this hypothesis. The results suggest that motor activity can retroactively interfere with the effectiveness of verbal information, as shown by later performace when no KR is furnished.

References

AMMONS, R. B. Effects of knowledge of performance: A survey and tentative theoretical formulation. *J. gen. Psychol.*, 1956, **54**, 279–299.

BILODEAU, E. A., & BILODEAU, I. M. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.

BOURNE, L. E., JR. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.

GOLDSTEIN, M., & RITTENHOUSE, C. H. Knowledge of results in the acquisition and transfer of a gunnery skill. *J. exp. Psychol.*, 1954, **48**, 187–196.

GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

GREENSPOON, J., & FOREMAN, S. Effect of delay of knowledge of results on learning a motor task. *J. exp. Psychol.*, 1956, **51**, 226–228.

HOUSTON, R. C. The function of knowledge of results in learning a complex motor skill. Unpublished master's thesis, Northwestern Univer., 1947.

JENKINS, W. O., & STANLEY, J. C. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.

JOHNSON, J. C. The effect of amount and timing of knowledge of results upon learning a motor skill. Paper read at Southeastern Psychol. Ass., April, Atlanta, 1956.

MCGUIGAN, F. J. Delay of knowledge of results: A problem in design. *Psychol. Rep.*, 1959, **5**, 241–243.

MCGUIGAN, F. J., CROCKETT, F., & BOLTON, C. The effect of knowledge of results before and after a response. *J. gen. Psychol.* (In press).

SALTZMAN, I. J., KANFER, F. H., & GREENSPOON, J. Delay of reward and human motor learning. *Psychol. Rep.*, 1955, **1**, 139–142.

(Received August 15, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

THE EFFECT OF DIFFERENTIAL ONSET TIME ON THE CONDITIONED RESPONSE STRENGTH TO ELEMENTS OF A STIMULUS COMPLEX[1]

DELOS D. WICKENS, ROBERT S. GEHMAN, AND SHIRLEY N. SULLIVAN

Ohio State University

In the typical learning situation in which some arbitrarily selected stimulus is temporally associated with reinforcement, it is inevitable that there will be many other stimuli which are likewise present when this same reinforcement occurs. In most psychological research the experimenter, especially the S–R theorist, is not interested in any increments in response strength which accrue to these unavoidably concurrent stimuli; he is interested in the accrual of response strength to his arbitrarily selected cue stimulus. Furthermore, when the results of the experiment are discussed, there is seldom a mention of these extra stimuli, and so there is some implication that it is only the *E*-designated cue stimulus which acquires response strength during the training.

The usual statement of the law of reinforcement, which simply stresses the relationship between the stimuli which are present, the response which is made, and the occurrence of reinforcement, would imply that these stimuli would also acquire habit strength. Furthermore, the reinforcement statement itself does not imply that the habit loadings acquired by these extra stimuli would differ from those acquired by the cue stimulus.

It seems most unlikely, however, that such would be the case. Indeed, the perceptually oriented behavior theorist is likely to insist that equality of effectiveness among stimuli would not occur. Tolman (1948), for example, has in his system a principle, essentially a postulate, that would lead to the prediction of inequality of effectiveness among stimuli. It is the concept of "searching for the stimulus." It implies that, in a complex environment, the organism will sort or classify the stimuli into those which are significant and those which are not.

In his *Principles of Behavior*, Hull (1943, Ch. 12) briefly considered this problem. He, too, assumed that inequality of habit loadings would accrue to the stimuli in a complex environment, and he offered some conjectures about factors which might be responsible for these inequalities. One of the factors he suggested was that of the time intervening between onset of the stimulus in question and occurrence of reinforcement. The present experiment deals with an investigation of this variable.

There is a considerable body of data (Kimble, 1947; Moeller, 1954; Reynolds, 1954; White & Schlosberg, 1952) which indicate that as the time between onset of the CS and the UCS increases from simultaneity to an interval in the neighborhood of 500 msec., the efficiency of conditioning increases. With further increase in time interval there is a decline in efficiency. It is possible to assume, therefore, that when more than one stimulus is present at the time reinforcement occurs, the response strength which accrues to each stimulus will be a

[1] This research was supported by the National Science Foundation under Grant G2248.

function of the length of time between its onset and the onset of reinforcement.

As a background for the present research, a brief description of an M.A. thesis supervised by the senior author is relevant (Lacey, 1952). The *Ss* were conditioned to a complex stimulus consisting of a light and tone, both of which overlapped and terminated with a shock. The behavior measured was the GSR. One of these stimuli came on 2000 msec. before shock, the other 500 msec. Following training, *Ss* were tested to one or the other element of the complex stimulus. Control groups conditioned to a single stimulus at the 2000- or 500-msec. intervals were also run. These control groups produced the expected finding since the response to the 500-msec. stimulus was greater than the response to the 2000-msec. stimulus during the extinction trials. The trend was reversed in the experimental groups, where a larger response was made to the element with 2000-msec. onset than to the element with the 500-msec. onset.

To account for the results, it may be assumed that a form of sensory conditioning can occur between the elements which compose the CS. During the presentation of the complex and the UCS, the complex not only acquires the capacity to evoke the GSR, but also one element of the complex acquires the capacity to evoke the "sensory effects" of the other element. When any element is presented alone, the magnitude of the GSR will be determined by how faithfully this element can produce the "sensory effects" of the other stimlus, and by combining this acquired effect with its own effect, approximate the sensory characteristics evoked by the complex itself.

One further assumption is required, and this is that something like the usual CS-UCS onset interval is applicable to the conditioning that accrues between the two elements of the complex. This implies that the shorter stimulus of the pair is placed in a backward relationship so far as the sensory conditioning is concerned. Such a relationship has been shown to be relatively ineffective in finger withdrawal conditioning (Spooner & Kellogg, 1947) and also in sensory preconditioning (Silver & Meyer, 1954). In a very general way this interpretation handles the Lacey data, for it implies that the long stimulus alone is more capable of evoking the stimulus pattern of the complex than is the short stimulus alone. As a result of stimulus generalization, different magnitudes of the GSR will be produced.

This interpretation implies that although the time relationship principle is applicable to the acquisition of response strength to elements of a complex, the point of reference for the measurement of time is the onset of the second CS rather than the onset of the UCS. As a first approximation, it may be assumed that the curve which has been obtained for time differences between onsets of a CS and UCS for a single CS will be applicable to the time differences of elements of a complex stimulus. It would therefore be predicted that as the time interval between the onset of the longer and shorter stimulus increased from simultaneity to an interval in the neighborhood of 500 msec., the longer element when presented alone would become increasingly effective. A decline in response strength would be associated with difference intervals greater than this. This statement applies to the case where the time of onset of the shorter stimulus with respect to reinforcement will be the same in all cases.

It is this prediction which was tested in the present experiment. A CS composed of two elements was employed. Following training to the

complex, different groups were extinguished to one or other element alone. The shorter or second stimulus of the complex preceded the shock UCS by the same time interval for all groups. The time by which the longer stimulus preceded the UCS was different for the various groups.

METHOD

Subjects.—The *Ss* were students in the elementary psychology course at Ohio State University who chose to volunteer for this experiment to meet part of the experimental time requirements of the course. A total of 598 *Ss* are represented in the data to be presented. The data of approximately 100 additional *Ss* were discarded for the following reasons: refusal to continue in the experiment because the shock was too strong, experimental errors, failure to respond to the shock on the last two conditioning trials, and resistance levels above 150,000 ohms. About 75 of the *Ss* were discarded for the last two reasons.

Apparatus.—The time between the onset of the two CS's and the UCS was controlled by an arrangement of wave form generators (Type 162), pulse generators (Type 163), and a power supply (Type 160 A) obtained from Tektronix, Inc. The error in this equipment is less than 1 msec. The timing was calibrated by means of a Tektronix Type 531 Oscilloscope. One CS was a 1000-cycle tone produced by a Jackson Oscillator. This tone was about 40 db (re .0002 dyne/cm.^{2}) and was set to match the psychological effectiveness of the visual stimulus. The matching was done by averaging judgments of several graduate students. This visual stimulus was the illumination of a 2-w. light concealed behind a ruby panel glass. The tone was fed into earphones; the light appeared at eye level, about 6 ft. directly in front of *S*. The shock was produced by an AC power pack. The shock electrodes were two pennies fixed on plexiglass clamps, one being placed on the palm of the hand and the other on the pad of the middle finger.

The GSR was recorded by means of the Fels Derhmometer, Serial 130. The instrument measures *S*'s resistance in ohms at any moment in time. At the sensitivity setting employed for most *Ss* it could be read accurately to units of 200 ohms. A small porportion of the *Ss* were run at a sensitivity setting where the meter units were 400 ohms. The *Ss* were usually those with relatively high resistance levels and who gave large responses. The recording electrodes were attached to the palm and back of the left hand. An electrode jelly of agar agar and zinc sulphate, made according to the formula recommended by the Yellow Springs Instrument Company, was used between the hand and the zinc electrodes.

The *S* sat in a dental chair in a 6 × 5 ft. semisoundproof room. It was dimly illuminated by a 7.5-w. purple bulb. An exhaust fan generated a masking noise, and together with the earphones and soundproofing prevented *S* from hearing the operating noises of the equipment in the experimental room.

The GSR was recorded visually. The resistance level at the time the stimulus was presented and also the minimum value immediately thereafter were recorded. These resistance scores were converted to reciprocals, multiplied by 10^4, and the difference between the two scores was obtained. The common log of this difference was used as *S*'s score on that trial.

Conditions and groups.—The experiment called for the use of 26 groups of *Ss*. Seventeen of these groups were conditioned to the stimulus complex of light and tone, and nine were conditioned to the tone alone. The two elements of the complex differed from each other in their duration prior to the onset of the shock. The shorter of the two stimuli began 500-msec. before shock and terminated simultaneously with it. The duration of the shock was 100 msec. The duration of the longer stimulus, which was coterminous with the short stimulus and the shock, varied from group to group. The schedule of these times is presented in Table 1. Following 10 reinforced trials to the complex of light and tone, five nonreinforced stimulations to one element of the complex were given. One group of *Ss* at each time interval was tested to the longer element of the complex, and another group was tested to the shorter element. The arrangement of the stimuli was counterbalanced for the groups during training so that the stimulus employed during the extinction series was always the tone. Hereafter, groups extinguished to the longer stimulus will be referred to as the Long Stimulus Groups, and those extinguished to the shorter stimulus will be called the Short Stimulus Groups. For each Long Stimulus Group of the complex, a single stimulus (*tone* the CS) Control Group was run. These

TABLE 1

THE TEMPORAL CHARACTERISTICS OF THE STIMULI FOR THE VARIOUS GROUPS OF THE EXPERIMENT

Exper. Groups	Time by Which Elements of the Complex Preceded UCS Onset		Time by Which Onset of CS_1 Preceded Onset of CS_2	Controls: Time by Which CS Onset Preceded UCS Onset
	Short	Long		
1	500	500	0	500
2	500	600	100	600
3	500	700	200	700
4	500	800	300	800
5	500	920	420	920
5	500	1040	540	1040
7	500	1550	1050	1550
8	500	2500	2000	2500
9	500	4600	4100	4600

groups, of course, matched the experimental Long Stimulus Groups insofar as the time of onset of the tone with respect to the time of onset of the shock was concerned.

Procedure.—After *Ss* were seated in the experimental cubicle, the following instructions were given to them:

"This is a conditioning experiment. First I want to put these electrodes on your left hand. Though I call them electrodes, you will get no shock from them; they simply record your response. You will feel nothing in this hand; just relax it and don't move it about much during the experiment. You will get a shock from the electrodes which I will place on the other hand. I'll begin the experiment by testing for the strength of the shock. I'll start with a weak one and increase it to the level we use in the experiment. The door will be open at this time so you can call out and let me know if it is too strong.

"After this is completed, I will come back in and put some earphones on you. During the experiment you will hear a tone in the earphones (and also this light will come on). This is the conditioned stimulus and it will be followed by the shock which is the unconditioned stimulus. All you have to do in the experiment is to keep your eyes open so that you will know when the light comes on. Of course you'll hear the tone through the earphones."

Three shocks alone of successively increasing intensity were presented, the final stimulation being 1.0 ma. at 220-v. This shock level was used for the first conditioning trial but during the succeeding five trials the shock was set up .2 ma. per trial. From the seventh trial on, the level used was 2.0 ma. at the same voltage. These values are based on a 60,000 ohm resistance across the electrodes.

Immediately after the shock was adjusted to the initial level, *E* entered the experimental cubicle, placed earphones on *S*, turned on the ventilation fan, and returned to the control room. The training consisted of 10 reinforced trials followed by 5 extinction trials. The intertrial intervals varied in a prearranged randomized sequence from 30 to 75 sec. in steps of 15 sec. The Experimental Groups were conditioned to the stimulus complex of light and tone and extinguished to the tone alone.

The *Ss* were assigned to the 26 groups in a constrained random order. The constraints involved attempting to match sexes in the various groups and attempting to equalize the numbers of *Ss* in the various groups at the end of each week.

TABLE 2

MEANS AND VARIANCES FOR ALL GROUPS ON TRIAL 1

Time Interval	Long Stimulus		Short Stimulus		Control	
	Mean	Variance	Mean	Variance	Mean	Variance
500	2.10957	.26818	2.10957	.26818	2.19831	.48270
600	2.01806	.72159	2.19140	.43862	2.14396	.29681
700	2.17953	.33331	2.13585	.29453	2.16077	.47168
800	2.11822	.42495	2.11542	.49441	1.98097	.27000
920	2.18071	.50261	1.97836	.58574	1.94092	.58706
1040	1.89891	.40550	2.26455	.33416	1.88523	.28633
1550	1.93960	.55728	2.20605	.21586	2.17660	.69468
2500	2.14078	.50576	1.96032	.28761	2.02816	.36942
4600	1.87622	.48452	1.97056	.62318	1.75177	.81586

TABLE 3

MEAN RESPONSE MAGNITUDE IN MICROMHOS DURING EXTINCTION TRIALS 2 + 3 AND 3 + 4

Time Interval	Long Stimulus		Short Stimulus		Control	
	2+3	4+5	2+3	4+5	2+3	4+5
500	1.91778	1.54620	1.91778	1.54620	2.02177	1.79207
600	1.81157	1.48949	1.88982	1.85813	1.87696	1.67447
700	1.96306	1.80165	1.70157	1.75065	2.01764	1.86181
800	1.86591	1.68143	1.81742	1.69938	1.76869	1.70042
920	1.86349	1.64216	1.74519	1.52056	1.79391	1.71225
1040	1.71594	1.54971	2.04583	1.91392	1.58064	1.52935
1550	1.86638	1.60783	1.90789	1.68706	2.00292	1.82222
2500	1.82999	1.60996	1.72277	1.55564	1.84668	1.77242
4600	1.60503	1.44364	1.86248	1.66044	1.40178	1.07148

RESULTS

The results are presented in Tables 2 and 3. Table 2 shows the mean and variance for all groups on Trial 1 of the extinction series. Table 3 shows the mean for the paired Trials 2 and 3 and the paired Trials 4 and 5. An examination of Trial 1 data for the Long Stimulus Groups indicates that there is a general tendency for the response magnitude to increase from the 600- to the 920-msec. group. Thereupon the response strength drops abruptly only to increase once again through the 2600-msec. group, and to decline in the 4600-msec. group. The Control means, however, show a general decline from the 600- through the 1040-msec. groups, and then they, too, rise for the middle groups and fall at the longest time interval. The means for the Short Stimulus Groups show a trend which is very nearly a perfect inverse to that shown by the Long Stimulus Groups. An examination of the data for the other extinction trials indicates that these same trends existed for all major groups throughout the extinction period.

The method of least squares was used to obtain a curve of best fit for each major group of the experiment. The 500-msec. interval was omitted from this analysis because this is the group in which the elements of the component came on simultaneously, and where its score could not be appropriately assigned to the Long or Short Stimulus Groups.[2]

The following equations were obtained:

$$\text{Long} \quad y = 1.89868 - .00361478x + .000130751x^2 + .00000172787x^3 - .000000016439x^4$$
$$\text{Short} \quad y = 2.23409 + .0026430x - .000112717x^2 - .00000198036x^3 + .000000017034x^4$$
$$\text{Control} \quad y = 2.18589 + .01043164x - .000117574x^2 - .00000497162x^3 + .000000034681x^4$$

where $x_a = \frac{1}{20}(t_a - \hat{t})$, with $\hat{t} = 1620$,

and t represents the time by which the long stimulus precedes the reinforcement. Each of the coefficients was tested against its standard error to determine whether its value differed significantly from zero. With the exception of one coefficient they were

[2] The authors wish to express their appreciation to D. Ranson Whitney and Lydia Kinzer for their supervision and performance of the statistical analysis.

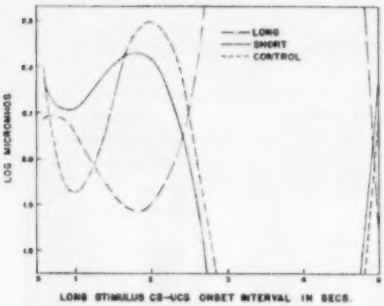

FIG. 1. The curves of best fit for the Long Stimulus, Short Stimulus, and Control Groups. The curves are interrupted in the range of 2700 to 4300 msec. because of their extremely high or low values.

all significant at the .01 level or better. The exception was the x^2 term for the Short Stimulus Group which was significant at the .03 level. The three curves generated by these functions are presented in Fig. 1.

DISCUSSION

Intervals Below 1100 Msec.

The control and long stimulus groups.—The most pertinent data for evaluating the stimulus generalization-sensory conditioning hypothesis are those for the 600- through 1040-msec. groups. The hypothesis assumes that the point of origin of the time relation curve is the onset of the short stimulus rather than the onset of the shock. Thus this time range is from 100 to 540 msec. prior to the hypothesized point of origin of the curve. Therefore, it would be predicted that the effectiveness of the Long Stimulus would increase during the major portion of this time range and then begin to decrease. This interpretation is based on an analogy to the Stimulus Reception curve as derived from studies with a single conditioned stimulus (Hull, 1952). This Stimulus Reception curve is thought to decline monotonically after 450 msec., but since both the Control and the Long Stimulus Groups evidence an unexpected rise in the neighborhood of 2000 msec., the initial evaluation of the hypothesis will be concerned only with the groups with a time interval below 1100 msec.

For the Long Stimulus Groups, the significant x and x^2 terms of the equation are particularly relevant to the evaluation of the hypothesis. Since, according to the conversion equation, values of x associated with times below 1620 msec. will be negative, and since the sign of the x term is negative, the slope term is positive. This is equivalent to stating that the curve rises for small values of t, and as the x^2 term, which is opposite in sign, begins to contribute more, the trend is reversed. The reversal occurs at approximately 800 msec. Such a trend is similar to that obtained for a single stimulus for time intervals from 100 to 600 msec.

For the Control Groups, however, the significant x term is opposite in sign, which means that the curve is falling from 600 msec. out to the neighborhood of 1000 msec. This trend conforms with the results of other *Es* (Moeller, 1954; White & Schlosberg, 1952).

The difference in performance of the Control and the Long Stimulus Groups can readily be accounted for by the assumption that in one case, the point of origin of the Stimulus Reception curve is the UCS, and in the other case it is the second CS. Granting this assumption, then results for the Long Stimulus Group can be fairly well predicted by the sensory conditioning-stimulus generalization hypothesis. One discrepancy should be noted: In this restricted range, the maximum of the best fit curve occurs about 300 msec. after the presumed origin of the curve rather than at 450 msec., as was assumed by Hull (1952).

The long and short stimulus groups.—The coefficients of the terms of both the Long and Short Stimulus equations are all significantly greater than zero and the parallel terms of the two equations are opposite in sign. Therefore, the two curves differ significantly from each other.

An inspection of the coefficients of the paired successive terms of the two equations shows them to be remarkably close to each other in value. This means that the curves, as Fig. 2 indicates, are almost perfect mirror images of each other. It is as if a certain magnitude of response could be conditioned at any given time interval, and, because of the particular time relations for that group, a certain amount of this response strength is controlled by one element of the complex, and control of the remainder is alloted to the other element. Thus, if the magnitude of the responses to the long and short stimulus for each time interval are added, and a plot is made of these values against time of onset of the long stimulus, these points fall very close to a straight line with a small negative slope.

These results seem, at least in a general way, to be consonant with the sensory conditioning hypothesis which gave rise to the experiment. The hypothesis assumes that the capacity to evoke the CR that is associated with either element of the complex is a function of the extent to which the element alone can reproduce the CS of the complex itself. If the time relations are favorable to produce sensory conditioning for one element, the very factors which make them favorable to this element may make them unfavorable to the other element.

Intervals Greater than 1100 Msec.

All three curves are characterized by significant x^3 and x^4 coefficients which are responsible for the complexity of the curves at the longer time intervals. For all three groups the trend in the curves in the range of 3000 to 4500 msec. does not seem to make good sense psychologically, for the predicted responses in this range are either extremely high or extremely low. This prediction undoubtedly results from the limited number of data points in this temporal region.

The scarcity of data points at the longer time intervals also results in some strange predictions for times greater than 4500 msec. The direction of the curves for the Long and Short Stimulus Groups fit expectations. One could predict that if a stimulus had an extremely long onset before reinforcement it would neither become an effective stimulus itself nor have an interaction with a stimulus which preceded reinforcement by a short time interval. The curve shows the Short Stimulus to be rising and the Long Stimulus to be falling. It should be noted, however, that the magnitude of response to the Short Stimulus when extrapolated to longer time intervals is unrealistically high.

The curve for the Control Group is clearly contradictory to expectations at longer time intervals, for extrapolation indicates that extremely large responses will occur with increased intervals of time. Actually, the psychological theory which gave rise to this experiment would have expected that the equation for this curve would be essentially the same as the equation for the Long Stimulus Group, but displaced 500 msec. to the left.

It seems likely, however, that the trend of the results in the range of 1200 to 2500 msec. should be taken more seriously. Despite the fact that the rise for the Control curve in this interval is contrary to other experimental findings (Moeller, 1954; White & Schlosberg, 1952), it is consistent with the other groups in this experiment. The Short Stimulus Group also shows a rise and decline in this region, while the Long Stimulus Group shows just the opposite trend. It would appear that an adequate explanation for the rise and fall in the Control curve should encompass the similar trend for the Short Stimulus curve and the reverse trend for the Long Stimulus curve.

In general, the derived curves seem to make very good sense psychologically at the shortest time intervals, are somewhat convincing at moderately long time intervals, but are clearly exaggerations for the longest time intervals.

The complexity of the obtained curves together with the relative simplicity of the experiment suggest that a more molecular explanation than that which gave rise to the experiment may eventually become possible. Such an explanation might be based upon our knowledge of action of the reticular formation and its relationship to behavioral data (Lindsley, 1958). This explanation might use the known effect of reticular activity on cortical activity (Moruzzi & Magoun, 1949); the influence of cortical activity on reticular activity (Adey, Segundo, & Livingston, 1957) and the influence of the reticular formation on the GSR (Wang, Stein, & Brown, 1956). These researches suggest that complex interacting relationships may be obtained as a consequence of slight temporal differences in the occurrence of stimulation.

Insofar as the broader problem of this research is concerned, the problem of predicting responses to elements of a complex from principles established in single stimulus studies, the results are ambiguous. They suggest that—at least for the GSR—the time relation curve established in single stimulus studies has some predictive value, but only if the point of origin of this curve is the other conditioned stimulus. It is obvious that the design of the present experiment makes it impossible to determine whether this effect is attributable to sensory conditioning directly, or to secondary reinforcing characteristics acquired by the second stimulus.

Summary

An experiment was undertaken to test the hypothesis that the curves showing the relationship between strength of conditioning and time of onset of the CS with respect to the UCS can be used to predict the strength of conditioning to elements of a stimulus complex.

Sixteen groups of *Ss* were conditioned to a complex CS consisting of a light and tone. One of these stimuli preceded the shock by 500 msec. for all groups. The time by which the other stimulus preceded shock ranged from 500 to 4600 msec. The response measured was the GSR. Following training, half the groups were tested to the short stimulus alone and half to the long stimulus alone. Control groups were conditioned to a single stimulus at the same time intervals as those employed for the long element of the complex CS.

When the response strength was plotted against the time by which the long stimulus preceded the UCS, functions whose coefficients were significant to the fourth degree were obtained. The curves for the Long and Short Stimulus groups of the complex were very close to being mirror images of each other. The function for the control group had coefficients which were opposite in sign to the function for the Long Stimulus of the complex.

In general, the data suggest that the time relation curve can crudely predict the response strength to an element of a complex, provided the point of origin of the curve is the onset of the second CS rather than the onset of the reinforcement. It is suggested, however, that a molecular explanation may be more adequate for handling all the data of the experiment.

References

Adey, W. R., Segundo, J. P., & Livingston, R. B. Corticifugal influences on intrinsic brain stem conduction in cat and monkey. *J. Neurophysiol.*, 1957, **20**, 1–16.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

Kimble, G. A. Conditioning as a function of the time between conditioned and unconditioned stimuli. *J. exp. Psychol.*, 1947, **37**, 1–15.

Lacey, R. J. Individual habit loadings as a function of time in complex stimuli. Unpublished master's thesis, Ohio State Univer., 1952.

Lindsley, D. B. The reticular system and perceptual discrimination. In *International symposium: Reticular formation of the brain.* Boston: Little, Brown, 1958.

Moruzzi, G., & Magoun, H. W. Brain stem reticular formation and activation of the EEG. *EEG clin. Neurophysiol.*, 1949, **1**, 455–473.

Moeller, G. The CS-UCS interval in GSR conditioning. *J. exp. Psychol.*, 1954, **48**, 162–166.

Reynolds, B. The acquisition of a trace conditioned response as a function of the

magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, **35**, 15–30.

SILVER, C. A., & MEYER, D. R. Temporal factors in sensory preconditioning. *J. comp. physiol. Psychol.*, 1954, **47**, 57–59.

SPOONER, A., & KELLOGG, W. N. The backward conditioning curve. *Amer. J. Psychol.*, 1947, **60**, 321–334.

TOLMAN, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, **55**, 189–208.

WANG, G. H., STEIN, P., & BROWN, V. W. Brainstem reticular system and galvanic skin reflex in acute decerebrate cats. *J. Neurophysiol.*, 1956, **19**, 350–355.

WHITE, C. T., & SCHLOSBERG, H. Degree of conditioning of the GSR as a function of the period of delay. *J. exp. Psychol.*, 1952, **43**, 357–362.

(Received August 18, 1958)

Journal of Experimental Psychology
Vol. 58, No. 1, 1959

SUPPLEMENTARY REPORT: ACQUISITION, EXTINCTION, AND COUNTERCONDITIONING WITH DIFFERENT VERBAL REINFORCEMENT COMBINATIONS[1]

ELSIE L. FERGUSON AND ARNOLD H. BUSS

University of Pittsburgh

A previous study (Buss, Braden, Orgel, & Buss, 1956) on verbal reinforcement combinations with psychiatric patients found that (*a*) the Right-Nothing (R-N) combination led to slower acquisition than the Right-Wrong (R-W) and Nothing-Wrong (N-W) combinations, and (*b*) extinction after the verbal reinforcement combinations was extremely slow. The present experiment extends the previous study, using normal *Ss*, expanding the extinction series, and adding counterconditioning.

Method.—The procedure and stimuli were the same as in the previous study. College men and women were randomly divided into three groups ($N = 20$ each) which received one of three reinforcement combinations during the 15-trial acquisition series: R-W, N-W, or R-N. Then each group was randomly divided in half, with 10 *Ss* having 100 extinction trials and 10 *Ss* having 100 counterconditioning trials.

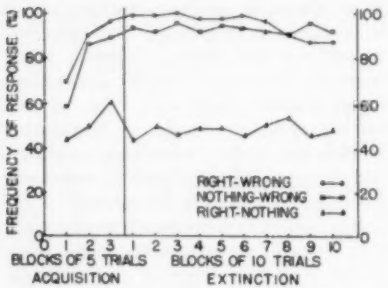

FIG. 1. Acquisition and extinction with different verbal reinforcement combinations.

Results.—The acquisition and extinction data are shown in Fig. 1, in which is shown the percentage frequency of the reinforced response. The R-W and N-W acquisition curves rise more steeply than the R-N curve. Despite random assignment of *Ss*, there were initial differences in response level. Therefore, difference scores (last 10 trials minus first 5 trials) were used in the analysis of variance of the acquisition data. The between-groups F was 17.88, $P < .001$. The frequency of the reinforced response in Group R-N was significantly lower than that of Group R-W ($t = 5.26$, $P < .001$) and Group N-W ($t = 5.06$, $P < .001$). The comparison of R-W with N-W frequencies gave $t < 1.0$. During extinction there was virtually no drop in level of responding, the curves being almost flat.

The counterconditioning data are presented in Fig. 2. The curves to the left represent the frequency of responses nonreinforced during acquisition; the curves to the right represent the same response, which was reinforced during counterconditioning. Groups R-W and N-W countercondition quickly and reach an asymptote of approximately 95% responding. Group R-N counterconditions more slowly, reaching an asymptote of approximately 70% responding. The between-groups F was 5.25 ($P < .05$), indicating the significantly poorer counterconditioning of Group R-N.

Discussion.—The results with normals are very similar to the results of the previous study with psychiatric patients. As in the previous study, R-N led to slower acquisition than the other two reinforcement combinations and there was extremely slow extinction. The relative absence of extinction cannot be attributed to insufficient trials,

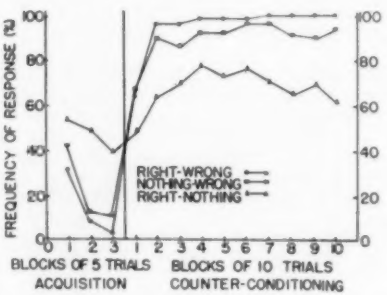

FIG. 2. Acquisition and counterconditioning with different verbal reinforcement combinations.

[1] Adapted from an M.S. thesis by the first author under the direction of the second author.

since in the present experiment the ratio of extinction trials to acquisition trials exceeded 6:1. In this situation there is evidently no appreciable extinction when verbal reinforcers are omitted.

Group R-N counterconditioned slower than Groups R-W and N-W. This finding, taken together with the acquisition results, is consistent with the hypothesis proposed in the previous study: *Right* is a weaker positive reinforcer than *Wrong* is a negative reinforcer, and *Nothing* is a nonreinforcer.

Reference

Buss, A. H., Braden, W., Orgel, A., & Buss, Edith H. Acquisition and extinction with different verbal reinforcement combinations. *J. exp. Psychol.*, 1956, **52**, 288–295.

(Received December 10, 1958)

JOURNAL OF EDUCATIONAL PSYCHOLOGY

This journal is now published by the American Psychological Association. In 1958 it became a bimonthly with issues in February, April, June, August, October, and December. Contents include articles on problems of teaching, learning, and the measurement of psychological development.

All back issues and subscriptions up to and including the May 1957 issue are the property of Warwick and York, Inc., 10 East Centre Street, Baltimore 2, Maryland.

Subscription $8.00
(Foreign, $8.50)

Single Copies, $1.50

Direct new subscriptions and renewals to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
Publications Office
1333 Sixteenth Street, N. W.
Washington 6, D. C.

PSYCHOLOGICAL MONOGRAPHS

The monographs listed below are still in print—but for how long?*

ADELBERT AMES, JR. *Visual Perception and the Rotating Trapezoidal Window.* No. 324. *$1.00*

JANE SHOHL. *Effects of Exposure to Sound on Discrimination Performance in the Rat.* No. 329. *$1.00*

PHILIP WORCHEL. *Space Perception and Orientation in the Blind.* No. 332. *$1.00*

P. R. WENDT. *Development of an Eye Camera for Use with Motion Pictures.* No. 339. *$1.00*

MONCRIEFF SMITH and EDNA A. WILSON. *A Model of the Auditory Threshold and Its Application to the Problem of the Multiple Observer.* No. 359. *$1.50*

M. R. ROSENZWEIG and W. A. ROSENBLITH. *Responses to Auditory Stimuli at the Cochlea and at the Auditory Cortex.* No. 363. *$1.00*

M. B. JONES. *An Experimental Study of Extinction.* No. 369. *$1.00*

G. F. RABIDEAU. *Differences in Visual Acuity. Measurements Obtained with Different Types of Targets.* No. 395. *$1.00*

---

* A selection from Volumes 65–69. (Write for complete list of in-print items.)

Order *now* from: **American Psychological Association**
1333 Sixteenth Street, N. W.
Washington 6, D. C.